DOCUMENT RESUME

ED 179 576 | **TM 009 857**

AUTHOR	Plattor, Emma; And Others
TITLE	Test Development for Assessing Achievement in Listening and Speaking.
INSTITUTION	Alberta Dept. of Education, Edmonton. Minister's Advisory Committee on Student Achievement.
SPONS AGENCY	Alberta Dept. of Education, Edmonton. Planning and Research Branch.
PUB DATE	Mar 79
NOTE	412p.
EDRS PRICE	MF01/PC17 Plus Postage.
DESCRIPTORS	Achievement Tests; Criterion Referenced Tests; *Educational Assessment; *Educational Objectives; Elementary Secondary Education; Factor Structure; Foreign Countries; Item Analysis; Listening Comprehension; Listening Skills; *Listening Tests; Speech Communication; Speech Skills; *Speech Tests; Student Attitudes; *Test Construction; Test Reliability; Test Validity
IDENTIFIERS	Canada

ABSTRACT

Pilot, field, and final tests in listening and in speaking knowledge, skills, and attitudes for Alberta (Canada) students in grades 3, 6, 9, and 12 were constructed, using the following steps: (1) selection and validation of high-priority measurable objectives; (2) review of existing tests and related literature; (3) pilot test construction, administration and data analysis; (4) construction, administration, and data analysis of a field test, based on pilot test findings; (5) final test construction; and (6) technical report writing. All items were constructed by the investigators because no usable items or tests were located during the literature review. Attitude items were excluded from the final test because they were unrelated to ability measures. Statistical analyses were performed on the pilot and field versions to determine item difficulty, item discrimination, reliability, and validity of multiple choice options. Factor analyses were also performed. The field test, with minor revisions, was recommended as a final test for large groups. Because criterion referenced testing in listening and speaking is relatively novel, however, the conclusions and recommendations made were tenuous. (The appendix contains the initial, pilot, field, and final test objectives for each grade). (CP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED179576

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Assessing Achievement in Listening and Speaking

MARCH, 1979

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
W. P. Eddy

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Alberta
Minister's Advisory Committee
on Student Achievement

CS 009 857

TEST DEVELOPMENT FOR ASSESSING ACHIEVEMENT IN LISTENING AND SPEAKING

A study Conducted for

The Minister's Advisory Committee on Student Achievement

(MACOSA)

by a

RESEARCH TEAM OF THE UNIVERSITY OF CALGARY

Emma Plattor - Faculty of Education
Department of Curriculum and Instruction
W. R. Unruh - Faculty of Education
Department of Educational Psychology
Laura Muir - Faculty of Education
Department of Curriculum and Instruction
K. D. Loose - Faculty of Continuing Education

through a contract with

ALBERTA EDUCATION

(Planning and Research Branch)

Project Completion Date: October, 1978

ABSTRACT

The purpose of this study was the development of criterion-referenced test instruments designed to assess Alberta (Canada) students' achievement of selected listening and speaking objectives at the Grades 3, 6, 9 and 12 levels. The principal investigators worked under the supervision of a Steering Committee appointed by the Minister's Advisory Committee on Student Achievement (MACOSA). The project was begun in February, 1978, and was completed in October, 1978.

Pilot, Field, and Final Test instruments in listening and in speaking at each of Grades 3, 6, 9, and 12 were constructed during the study, which consisted of the following phases: (1) selection, validation, and prioritization of measurable objectives in listening, speaking, and communication attitudes appropriate to each of Grades 3, 6, 9, and 12; (2) review of the literature in communication, listening, and speaking generally and measurement in listening, speaking, and communication attitudes specifically; (3) construction of Pilot Test instruments in listening, speaking, and communication attitudes for each of Grades 3, 6, 9, and 12; (4) administration of these Pilot Tests to samples of Alberta students at each grade level; (5) analysis of the data obtained from these Pilot Test administrations; (6) construction of Field Test instruments in listening, speaking, and communication attitudes at each grade level based on the findings of the Pilot Test data analyses; (7) administration of these Field Tests to samples of Alberta students at each grade level; (8) analysis of the data obtained from the Field Test administrations; (9) construction of Final Test instruments in listening and speaking for each grade level based on the results of the Field Test data analyses. (These analyses indicated that the scales used to measure attitudes in this study did not perform this function or appear to be related to the types of activity used in the Listening and Speaking Tests, so this instrument was dropped.); and (10) development of a Technical Report of the study, as well as a Condensed Report, and Abstract.

Knowledges, skills, and attitudes in listening and speaking were tested during the Pilot and Field Test phases. Since no usable tests or items in listening and speaking were located during the literature review, all items were constructed by the investigators for all of the Listening and Speaking

Tests. Two commercial Communication Attitude Tests were administered in the Pilot and Field Test phases but were not included in the Final Tests.

The Final Speaking Tests at each grade level consisted of two parts. Part 1, Speaking (Written), consisted of stimulus subtests designed to measure student knowledges in both speaking and listening (communication competency). Knowledge of selected techniques for oral reading (interpretation) was tested at all levels. At all grade levels, knowledge of organizational factors was also tested in this part of the Speaking Tests. Part 2, Speaking (Written), was designed to test student skills in oral reading (interpretation) and in spontaneous speech during an audience situation (a small group of 4-6 of the students' peers). At the Grade 3 level, this part of the Speaking Tests also contained an articulation subtest. Multiple choice items were used in Part 1, and Rating Scales were used in Part 2. The items used in both parts were based on domains into which the selected objectives had been classified as well as on the objectives themselves. Teacher's Manuals and Student Test Booklets were developed for administration at each grade level.

Pilot testing was carried out in intact classes in Calgary and Okotoks, Alberta. Field testing was done in intact classes in Didsbury and Red Deer, Alberta. The Pilot Test sample was the same for both the Listening and Speaking Test administrations. The Field Test sample was also the same for both the Listening and Speaking Test administrations.

Statistical analyses were performed on both the Pilot and Field Listening and Speaking Tests in order to determine: (1) if the options in the multiple choice items worked properly; (2) if the items were too difficult or too easy; (3) if the items discriminated among students; and (4) if the ratings and tests themselves were reliable. Factor analyses were also done. The Pilot and Field Tests were revised based on these statistical analyses. The Final Listening and Speaking Tests reflected these revisions.

Conclusions and recommendations based on the study were presented in the Technical Report.

FOREWORD

The Minister's Advisory Committee on Student Achievement (MACOSA) was established by ministerial order in October 1976 in response to growing concerns expressed by the public at large. government, labor, business, students and educators regarding the quality and standards of basic education in Alberta.

MACOSA commissioned a number of studies, primarily to provide basic information for a summary of current levels of achievement in Alberta and to provide baseline data for future assessment. These studies fell into three categories: (1) preliminary studies, (2) achievement studies, and (3) other studies.

This study, Test Development for Assessing Achievement in Listening and Speaking, was commissioned to design or adapt test for assessing the achievement of selected objectives in listening and speaking in grades 3, 6, 9 and 12.

This report, which represents the findings and conclusions of the researchers, was presented to MACOSA as information.

MACOSA LISTENING AND SPEAKING

STEERING COMMITTEE

Mr. Bernard Gommeringer (Chairman),
Language Arts Consultant, Alberta Education

Dr. Wilma Laing, Faculty of Education
University of Alberta

Dr. Clarence Rhodes, Consultant, Planning and Research
Branch, Alberta Education

Dr. William Washburn, Calgary Board of Education

Dr. Harry Sherk (ex officio), Associate Director,
Planning and Research Branch, Alberta Education

Consultant to the Committee:
Dr. Merv Thornton, Associate Director of Curriculum,
Alberta Education

TABLE OF CONTENTS

Page

Foreword i

MACOSA Listening and Speaking Steering Committee ii

Table of Contents iii

List of Tables x

List of Figures xvi

PART I: CONDENSED REPORT

- Introduction 1
- Organization of the Study (Work Plan) 3
- Review of Related Literature 5
- Theoretical Framework 6
- Design of the Study 13
  - Rationale and Preliminary Planning 13
  - Design Steps 17
  - Selection of Objectives 20
  - Item Development 21
- Pilot Test Administration and Data Analysis. 23
  - Sample 23
  - Procedure 23
  - Analysis of Pilot Test Data. 24
  - Rater Reliability 25
- Field Test Administration and Data Analysis 26
  - Sample 26
  - Procedures 26
  - Analysis of Field Test Data 26
    - Item Analysis 26
    - Descriptive Statistics 27
    - Correlational Analysis 28
    - Factor Analysis 29
- Conclusions and Recommendations 30

Table of Contents (Continued)

Page

Future Test Administration. 30

Future Testing and Date Analysis. 34

Future Test Development 35

Attutude Testing. 37

Reporting Test Results. 37

Teacher Training. 39

References 40

PART II: TECHNICAL REPORT. 41

PREFACE 42

1 INTRODUCTION TO THE STUDY. 45

Purposes of the Listening/Speaking Assessment 48

Theoretical Framework 49

Limitations of the Assessment 57

Organization of the Study (Work Plan) 57

Organization of the Report. 63

2 REVIEW OF THE LITERATURE 65

Functional Communication. 67

Listening 69

Speaking. 78

Attitude. 85

3 DESIGN AND IMPLEMENTATION OF THE STUDY 89

Overview. 89

Rationale 89

Design Steps. 96

Objectives 106

Listening/Speaking Knowledges 107

Listening/Speaking Skills 109

Attitudes 110

Domain Development for Pilot Testing. 111

Item Development. 115

Table of Contents (Continued)

Page

3 Listening: Development of the Pilot Tests 117

Language Types 117

Stimulus Subtests 119

Item Format and Selection 121

Administration Time 121

Administration Manual 127

Stimulus Materials 127

Student Test Booklets 127

Audio-Tape Recordings 132

Listening: Development of the Field Tests 132

Language Types 132

Stimulus Subtests 134

Item Format and Selection 136

Administration Time 136

Administration Manual 142

Stimulus Materials 142

Student Test Booklets 142

Audio-Tape Recordings 147

Speaking: Development of the Pilot Tests 148

Speaking (Written) 149

Speaking (Oral) 149

Stimulus Materials 152

Administration Time 163

Table of Contents (Continued)

Page

3 Administration Manual 164

Student Test Booklets 164

Attitude Measure 165

Speaking: Development of the Field Tests 165

Speaking Field Test Procedures 176

Administration Time 177

Listening and Speaking: Final Tests 177

4 DATA ANALYSIS AND DISCUSSION 179

The Pilot Tests 179

Sample 180

Pilot Test Administration 181

Pilot Listening Tests: Procedures 182

Pilot Speaking Tests: Procedures 183

General Procedures 183

Specific Procedures 184

Pilot Test Results 186

Listening 186

Speaking (Written) 192

Speaking (Rating Scales) 195

Rater Reliability 195

The Field Tests 196

Sample 197

Table of Contents (Continued)

Page

4 Field Test Administration 199
Item Analysis Data 204
  Listening 204
  Speaking (Written) 214
  Speaking (Ratings) 219
Descriptive Statistics 228
  Listening 228
  Speaking (Written) 233
Correlational Analysis by Stimulus Sub-Scales . . . 236
  Listening 236
  Speaking (Written) 241
General Correlational Analysis 247
  Grade 3 247
  Grade 6 249
  Grade 9 251
  Grade 12 253
Factor Analysis 255
  Listening 255
    Grade 3 255
    Grades 6 - 12 261
  Speaking 262
    Grade 3 262
    Grade 6 267

Table of Contents (Continued)

Page

4 Grade 9 270

Grade 12 282

Attitude 290

5 RECOMMENDATIONS AND CONCLUSIONS 293

6 REFERENCES 303

APPENDICES 311

A-1 Initial Listening/Speaking Knowledge Objectives . . 312

A-2 Initial Listening Skill Objectives 318

A-3 Initial Speaking Skill Objectives 328

A-4 Attitude Objectives 335

B-1 Initial, Pilot, Field, and Final Listening Test Objectives (Grades 3, 6, 9, 12) 337

B-2 Initial, Pilot, Field and Final Speaking Test Objectives (Grades 3, 6, 9, 12) 341

C-1 Final Listening Test Objectives - Grade 3 346

C-2 Final Listening Test Objectives - Grade 6 347

C-3 Final Listening Test Objectives - Grade 9 349

C-4 Final Listening Test Objectives - Grade 12 351

C-5 Final Speaking Test Objectives - Grade 3 353

C-6 Final Speaking Test Objectives - Grade 6 355

C-7 Final Speaking Test Objectives - Grade 9 357

C-8 Final Speaking Test Objectives - Grade 12 359

D-1 Final Listening Test Objectives/Item Matrix Grade 3 361

D-2 Final Listening Test Objectives/Item Matrix Grade 6 362

D-3 Final Listening Test Objectives/Item Matrix Grade 9 363

Table of Contents (Continued)

Page

APPENDICES

D-4 Final Listening Test Objectives/Item Matrix Grade 12 364

D-5 Final Speaking Test Objectives/Item Matrix Grade 3 . 365

D-6 Final Speaking Test Objectives/Item Matrix Grade 6 . 366

D-7 Final Speaking Test Objectives/Item Matrix Grade 9 . 367

D-8 Final Speaking Test Objectives/Item Matrix Grade 12 369

LIST OF TABLES

TABLE Page

3-1 Listening Pilot Test Items by Grade and Language Type 118

3-2 Listening Pilot Test Content by Stimulus Subtest, Stimulus Format, Language Type, Item Numbers, and Number and Percent of Items . . 120

3-3 Listening Pilot Test Blueprint, Grade 3 122

3-4 Listening Pilot Test Blueprint, Grade 6 123

3-5 Listening Pilot Test Blueprint, Grade 9 124

3-6 Listening Pilot Test Blueprint, Grade 12 125

3-7 Listening Pilot Test Objectives/Items Matrices, Grades 3,6,9,12 126

3-8 Listening Pilot Test Scripts and Sources, Grade 3 128

3-9 Listening Pilot Test Scripts and Sources, Grade 6 129

3-10 Listening Pilot Test Scripts and Sources, Grade 9 130

3-11 Listening Pilot Test Scripts and Sources, Grade 12 131

3-12 Listening Field Test Items by Grade and Language Type 132

3-13 Listening Field Test Content by Stimulus Subtest, Stimulus Format, Language Type, Item Number and Number and Percent of Items 135

3-14 Listening Field Test Blueprint, Grade 3 137

3-15 Listening Field Test Blueprint, Grade 6 138

3-16 Listening Field Test Blueprint, Grade 9 139

3-17 Listening Field Test Blueprint, Grade 12 140

TABLE Page

3-18 Listening Field Objectives/Item Matrices, Grades 3,6,9,12 141

3-19 Listening Field Test Scripts and Sources, Grade 3 143

3-20 Listening Field Test Scripts and Sources, Grade 6 144

3-21 Listening Field Test Scripts and Sources, Grade 9 145

3-22 Listening Field Test Scripts and Sources, Grade 12 146

3-23 Speaking Pilot Test Items by Grade and Language Type 152

3-24 Speaking Pilot Test Scripts and Sources, Grades 3,6,9,12 153

3-25 Speaking Pilot Test Content by Stimulus Subtest, Stimulus Format, Language Type, Item Numbers, and Number and Percent of Items 157

3-26 Speaking Pilot Test Blueprint, Grade 3 158

3-27 Speaking Pilot Test Blueprint, Grade 6 159

3-28 Speaking Pilot Test Blueprint, Grade 9 160

3-29 Speaking Pilot Test Blueprint, Grade 12 161

3-30 Speaking Pilot Test Objectives/ Items Matrix, Grades 3,6,9, 12 162

3-31 Speaking Field Test Items by Grade and Language Type 167

3-32 Speaking Field Test Scripts and Sources, Grades 3,6,9, 12 168

3-33 Speaking Field Test Content by Stimulus Subtest, Stimulus Format, Language Type, Item Number, and Number and Percent of Items. . . 170

TABLE Page

3-34 Speaking Field Test Blueprint, Grade 3 171

3-35 Speaking Field Test Blueprint, Grade 6 172

3-36 Speaking Field Test Blueprint, Grade 9 173

3-37 Speaking Field Test Blueprint, Grade 12 174

3-38 Speaking Field Test Objectives/ Items Matrices, Grades 3,6,9, 12 175

4-1 Item Analysis Data - Grade 3 Listening (Pilot Test) 187

4-2 Item Analysis Data - Grade 6 Listening (Pilot Test) 189

4-3 Item Analysis Data - Grade 9 Listening (Pilot Test) 190

4-4 Item Analysis Data - Grade 12 Listening (Pilot Test) 191

4-5 Means, Standard Deviations and Reliability Coefficients for Listening Pilot Tests 192

4-6 Item Analysis Data - Grade 3 Speaking (Written) (Pilot Test) 193

4-7 Item Analysis Data - Grade 6 Speaking (Written) (Pilot Test) 193

4-8 Item Analysis Data - Grade 9 Speaking (Written) (Pilot Test) 194

4-9 Item Analysis Data - Grade 12 Speaking (Written) (Pilot Test) 194

4-10 Means, Standard Deviations and Reliability Coefficients For Speaking Pilot Tests (Written) . . 195

4-11 Item Analysis Data - Grade 3 Listening Field Test 205

4-12 Item Analysis Data - Grade 6 Listening Field Test 207

TABLE Page

4-13 Item Analysis Data - Grade 9 Listening Field Test 209

4-14 Item Analysis Data - Grade 12 Listening Field Test 211

4-15 Item Analysis Data - Grade 3 Speaking Field Test (Written) 215

4-16 Item Analysis Data - Grade 6 Speaking Field Test (Written) 216

4-17 Item Analysis Data - Grade 9 Speaking Field Test (Written) 217

4-18 Item Analysis Data - Grade 12 Speaking Field Test (Written) 218

4-19 Item Analysis Data - Grade 3 Speaking Field Test 220

4-20 Item Analysis Data - Grade 6 Speaking Field Test 221

4-21 Item Analysis Data - Grade 9 Speaking Field Test 223

4-22 Item Analysis Data - Grade 12 Speaking Field Test 225

4-23 Statistics for Stimulus Sub-scales and Total Grade 3 Listening Field Test 229

4-24 Statistics for Stimulus Sub-scales and Total Grade 6 Listening Field Test 230

4-25 Statistics for Stimulus Sub-scales and Total Grade 9 Listening Field Test 231

4-26 Statistics for Stimulus Sub-scales and Total Grade 12 Listening Field Test 232

4-27 Statistics for Stimulus Sub-scales and Total Grade 3 Speaking Field Test (Written) . 234

TABLE Page

4-28 Statistics for Stimulus Sub-scales and Total Grade 6 Speaking Field Test (Written) . . 234

4-29 Statistics for Stimulus Sub-scales and Total Grade 9 Speaking Field Test (Written) . . 235

4-30 Statistics for Stimulus Sub-scales and Total Grade 12 Speaking Field Test (Written) . . 235

4-31 Correlational Analysis of Grade 3 Listening Field Test by Stimulus Sub-scale and Sex . . . 237

4-32 Correlational Analysis of Grade 6 Listening Field Test by Stimulus Sub-scale and Sex . . . 238

4-33 Correlational Analysis of Grade 9 Listening Field Test by Stimulus Sub-scale and Sex . . . 239

4-34 Correlational Analysis of Grade 12 Listening Field Test by Stimulus Sub-scale and Sex . . . 240

4-35 Correlational Analysis of Grade 3 Speaking Field Test (Written) by Stimulus Sub-scales and Sex 242

4-36 Correlational Analysis of Grade 6 Speaking Field Test (Written) by Stimulus Sub-scales and Sex 243

4-37 Correlational Analysis of Grade [illegible] Speaking Field Test (Written) by Stimulus Sub-scales and Sex 244

4-38 Correlational Analysis of Grade 12 Speaking Field Test (Written) by Stimulus Sub-scales and Sex 245

4-39 Correlation Matrix for Listening, Speaking and Attitude by Domain Sub-scales and Total Test - Grade 3 248

4-40 Correlational Matrix for Listening, Speaking and Attitude by Domain Sub-scales and Total Test - Grade 6 250

TABLE Page

4-41 Correlation Matrix for Listening, Speaking and Attitude by Domain Sub-scales and Total Test - Grade 9 252

4-42 Correlation Matrix for Listening, Speaking and Attitude by Domain Sub-scales and Total Test - Grade 12 254

4-43 Factor Structure for Grade 3 Listening Test . . . 256

4-44 Factor Labels and Comparison with A-priori Domains - Grade 3 Listening 258

4-45 Factor Structure for Grade 3 Speaking Test 263

4-46 Factor Labels and Comparison with A-priori Domains - Grade 3 Speaking 264

4-47 Factor Structure for Grade 6 Speaking Test 268

4-48 Factor Labels and Comparison with A-priori Domains - Grade 6 Speaking 271

4-49 Factor Structure for Grade 9 Speaking Test 275

4-50 Factor Labels and Comparison with A-priori Domains - Grade 9 Speaking 278

4-51 Factor Structure for Grade 12 Speaking Test . . . 283

4-52 Factor Labels and Comparison with A-priori Domains - Grade 12 Speaking 286

4-53 Means and Standard Deviations for PRCF and PRCA 291

LIST OF FIGURES

Page

FIGURE

1. Communication Process - Flow Chart 19, 98

TEST DEVELOPMENT FOR ASSESSING ACHIEVEMENT IN LISTENING AND SPEAKING

PART I

CONDENSED REPORT

INTRODUCTION

The Minister's Advisory Committee on Student Achievement (MACOSA) was established by Ministerial Order in October, 1976, to identify solid evidence of the level of achievement in Alberta schools (Alberta Education, 1977). The Committee's mandate included:

1. studying the elements of student achievement, including definition, purpose, principles, and assumptions underlying achievement;
2. consideration of ways and means to assess student achievement;
3. consideration of methods for maintaining or improving achievement; and
4. recommendations on these matters as based on findings.

In order to fulfill this mandate, MACOSA commissioned a number of studies in various achievement categories. On February 7, 1978, Alberta Education contracted the University of Calgary and specifically the principal investigators E. Plattor, W. Unruh, and L. Muir, professors in the Department of Curriculum and Instruction and of Educational Psychology, to design and/or adapt tests for assessing the achievement of selected objectives in listening and speaking for Grades 3, 6, 9, and 12.

The investigators worked under the supervision of a MACOSA-appointed Steering Committee to:

1. prepare lists of measurable listening/speaking objectives based on ones prioritized and validated by the Steering Committee in cooperation with the Curriculum Branch of Alberta Education and their Language Arts Coordinating Committee;
2. review related literature in listening/speaking;

3. construct an item pool for assessing selected listening/speaking objectives;
4. construct and administer pilot test instruments to a sample of students at Grades 3, 6, 9 and 12;
5. analyze the data obtained from the pilot tests and interpret these data in order to provide baseline information for the production of field test instruments;
6. construct and administer field test instruments to a sample of students at Grades 3, 6, 9 and 12;
7. analyze the data obtained from the field tests and interpret these data in order to provide baseline information for the production of final test instruments;
8. construct final test instruments in listening and speaking at the Grade 3, 6, 9, and 12 levels;
9. compile into a Technical Report all of the information obtained from the pilot and field testing together with recommendations.

A Prospectus entitled Assessing Achievement of Selected Objectives in Listening and Speaking was submitted by the principal investigators to Alberta Education on January 10, 1978. This was followed by a request from Alberta Education for a formal Proposal for the study. A Proposal, also entitled Assessing Achievement of Selected Objectives in Listening and Speaking, was submitted on February 1, 1978. This Proposal was accepted by Alberta Education, and a contract between Alberta Education and the University of Calgary for a project entitled MACOSA-Developing and Validating Achievement Tests in Speaking and Listening was finalized on February 7, 1978.

At Steering Committee meetings held on March 7, and April 6, 1978, a number of guidelines governing the study were agreed upon. The tests were to take a maximum of one hour of student time; technological aids--tape recordings, etc.--would be used where feasible; as much as possible, testing and observation would be conducted in natural settings; statistical analyses would make use of traditional procedures for evaluation of items; common stimuli would be used as much as possible across at least two grades so that developmental trends could be observed; when stimulus materials required the tape-recording of speaking activities, students and classroom teachers would be used in making the tapes; and, wherever possible, knowledges and attitudes would be tested by the use of pencil-and-paper tests, while skills would be evaluated directly.

The Research Team now proceeded to devise test instruments for the following purposes:

1. To provide for program evaluation through the collection of baseline data in listening and speaking.
2. To provide for initial diagnostic assessment of individual students based on the achievement of specific objectives. For this purpose, measures of individual achievement would be a secondary concern.
3. If possible, to provide teachers with models for their use in developing tests in listening/speaking based on specific objectives appropriate for their own students.

ORGANIZATION OF THE STUDY (WORK PLAN)

The organization of the study can be stated in terms of activities summarized as follows:

Activity 1: The selection and development of significant, measurable objectives related to knowledges, abilities, and skills. (See the Technical Report for a description of this process and the sources used.)

Activity 2: The selection of an Advisory Panel of experts to respond to items and objectives. (Because of time pressures and the need to expend funds for test administration, this Panel was not used. However, ongoing advice was solicited from colleagues and knowledgable teachers throughout the study.)

Activity 3: An extensive search of the literature for the purpose of locating objective statements, possible test items, and related research findings.

Activity 4: The development of Test Blueprints and Items/Objectives Matrices.

Activity 5: The construction of an item pool to test objectives which had been selected, validated, and prioritized by the Research Team and the Steering Committee. (No usable items were located in the literature search. Thus all items were constructed by the Research Team.)

Activity 6: The development and administration of Pilot Tests in listening, speaking and attitude toward communication. (These tests were conducted in May, 1978 in selected schools in Calgary, Okotoks, Airdrie and Springbank.)

Activity 7: The development of the Field Tests and Sampling Plan. Grades 3, 6 and 9 were tested in June, 1978. Because of testing difficulties in Grade 12 at this time of year, this portion of the sample was tested in September, 1978. The Field Tests were conducted in schools in Didsbury and Red Deer. Regular classroom teachers were used for administration of the Listening Field Tests, the Written portion of the Speaking Field Tests and the Attitude Tests. Certified substitute teachers were hired and trained by the Research Team to administer those portions of the test which required rating of individual pupils.

Activity 8: The development of a Data Analysis Plan, scoring procedures and a Reporting Plan.

Activity 9: The production of a Technical Report, a Condensed Report and the Final Test Instruments. Instruments. As a preliminary step in the production of the Technical Report and Final Test Instruments, an outline of each of these was presented to the Steering Committee for comment and approval.

REVIEW OF RELATED LITERATURE

A major task in this study involved a review of the literature in listening, speaking, and general communication in order to validate the Research Team's conception of behaviors and instructional domains in listening/speaking, as well as to analyze existing evaluation instruments.

Several research reports were commissioned by MACOSA prior to the inception of this study. The Marchak Report (1979) examined procedures and instruments used in assessing communication skills and provided a number of recommendations for the development of instruments designed to test listening and speaking. The Plattor Report (1977) examined selected research reports in the area of listening and also made recommendations for test development.

The literature review was begun with an extensive search of ERIC documents and those of Psychological Abstracts. Many curriculum guides, professional textbooks, Teacher's Guides accompanying textbook series in language arts areas, and student texts were examined in an effort to determine listening/speaking objectives and curriculum materials currently in use in North America and, where appropriate, in England. Materials concerned with listening, speaking, and communication that were published by professional educational associations (e.g.: National Council of Teachers of English, International Society for General Semantics, Speech Communication Association, etc.) were also examined. Published reviews of measurement devices as well as other published materials concerned with the assessment of communication, speaking, listening, and attitude instruments were reviewed.

Chapter 2 of the Technical Report summarizes the findings and conclusions reported in the literature. This chapter includes a discussion of functional communication, the elements of listening and speaking skills, guidelines for testing and test construction in these areas, and a discussion of attitudes towards communication events and how such attitudes might be evaluated. On the basis of this review a theoretical framework to guide the study was constructed.

THEORETICAL FRAMEWORK

Oral communication involves both the reception and expression of ideas and information through spoken language. Effective oral communication not only conveys intended messages and meanings; it is also appropriate to the communication situation, the audience, and the speaker's purpose. In a recently published book, Larson et. al. (1978) point out that functional communication refers to the skills, knowledge, and attitudes possessed by an individual, from which that individual's competence or effectiveness may be inferred. They further note that the levels of functional communication, above competence, imply degrees of communication effectiveness. Where competence involves minimal levels of communicative functioning, effectiveness involves communicating in such a way that certain desirable outcomes are enhanced or facilitated. However, a person may be an effective communicator in many different ways. As Larson et. al. point out:

> If another person understands the content of your communication or understands your unique personal characteristics, you may be effective in the sense that you have facilitated or enhanced accurate perception between you and the other. If you are able to change or modify another person's point of view on an issue, you may be considered effective in the sense that you have facilitated or influenced an outcome. . . . Many different outcomes are used both to conceptualize and to measure an individual's level of communication effectiveness. For the moment, the point is simply that communication effectiveness implies something more than, and different from, communication competence. (p. 3).

Larson et. al. also suggest that the difference between competence and effectiveness is paralleled generally by different measurement strategies. They note that competence is usually assessed through objective-referenced or criterion-referenced tests, the purpose of which is to determine whether or not a person demonstrates a specific ability or performs a particular task. The test score indicates whether or not the person taking the test responds correctly or adequately to the test items. Such tests are appropriate when examining the capability of students to perform each of a variety of communication operations. This type of assessment is different from norm-referenced testing, where the purpose is to determine how well the student performs certain communication operations. In such testing, the score indicates how well the student can demonstrate a skill level compared to how well others being assessed can demonstrate this skill level. Norm-referenced testing is not concerned with the presence or absence of a particular communication knowledge or skill. Since the purpose of the MACOSA project was to develop test instruments in listening/speaking that could provide for program evaluation through the collection of baseline data about general student functioning, as well as to provide for diagnostic assessment of individual students rather than for comparison among students, it appeared reasonable to develop criterion-referenced tests which would provide evidence relative to whether or not students demonstrated specific abilities--not how well they demonstrated such abilities in comparison with each other.

While the importance of oral communication is accepted by educators, considerably greater emphasis at all grade levels has been placed on the teaching of written reception and expression. As a result, little attention has been given to defining the objectives of listening/speaking within language arts programs at the elementary and secondary school levels. For example, Landry

(1969) points out that "Although research has consistently shown that pupils spend more time in listening than in any other language arts activity, a serious lack of programs which develop listening skills is evident in most elementary schools." Devine (1978), in a study of listening research and theory over the past fifty years, notes that "Listening . . . is not generally accepted as part of the standard school curriculum. Surveys of actual classroom practices indicate little time devoted to direct instruction in listening skills, even though the evidence indicates strongly that such instruction works."

Examination of curriculum handbooks and guides developed by various Alberta school divisions reveals that most include some oral communication objectives. While these vary in number and specificity, they are usually concerned with oral reception and expression in formal rather than in functional communication situations. Since emphasis is generally on the use of language to inform, other uses of language are often dealt with minimally or are totally neglected. The predominance of written expression in the curriculum is reflected in the fact that listening objectives appear to be mirror images of reading objectives and speaking objectives appear to be mirror images of writing objectives. Thus concern in listening is generally with those skills involved in comprehending written language read aloud rather than in understanding spoken language, and concern in speaking is generally with those skills involved in reporting and oral interpretation rather than with those skills involved in functional and interpersonal oral communication. Nevertheless, current curriculum development in Alberta places strong emphasis on the development of functional oral communication skills within a balanced language arts program. Thus the development of appropriate oral communication objectives and assessment devices is an urgent task in Alberta at this time.

Several of the recommendations contained in two reviews of language arts literature developed for MACOSA in 1977 emphasize the importance of developing objectives and assessment instruments that examine oral communication in functional situations for various audiences and purposes (Marchak, 1979; Plattor, 1977). These documents further reinforce the importance of assessment in classroom-type situations. Among these recommendations are the following:

1. Since oral reception and expression occur in situational contexts, it is most reasonable to measure these types of expression in such contexts. Therefore, controlled but natural testing situations should be used in assessment.
2. Reception and expression of spoken rather than written language should be a major concern in the construction of oral language assessment instruments.
3. Oral expression (speaking competence) should be tested in individual or small group situations.
4. Assessment of oral reception and expression should focus on the elements of spoken language that make it different from written language.
5. Since the communication of ideas and information involves both cognitive and affective aspects, both of these should be reflected in oral communication objectives and in assessment instruments designed to test oral communication.

In a recent paper (1978), Nancy Mead of the National Assessment of Educational Progress (NAEP) provides some additional recommendations for developing oral communication tests. (These recommendations are based on NAEP's experiences in attempting to develop listening tests.) They include:

1. Focus on skills which are unique and central to listening (and, by implication, to speaking). Mead notes that these skills involve responding to oral language, which is different from spoken language in that it tends to be nonlinear, incomplete, and redundant. However, she also notes that "It is less obvious how to deal with the nonverbal and interactive nature of oral communication in an assessment situation. Nonverbal signals tend to be subtle and individualistic and thus difficult to include in an

> assessment. Likewise, the give and take of normal speaking and listening are difficult to recreate in a test setting." (p. 10).

2. Use short, interesting, listening stimuli. As Mead notes, by focusing on topics and listening situations which are common to most students (for example, school activities, friendships, television), it is possible to provide stimuli that are both natural and interesting.
3. Consider extraneous factors which might contribute to item bias, e.g., the quality of the listening stimulus, the language level used, etc.

It is interesting to note the disclaimer following these recommendations:

> The recommendations discussed above are general and suggestive . . . . The recommendations emerge more from the subjective experiences of the NAEP/SCA pilot listening assessment project than from concrete empirical findings. These guidelines must be subjected to careful study and research . . . . (p. 14).

Nevertheless, the investigators in the MACOSA speaking/listening assessment development project used these as guidelines both in designing the tests and in drawing conclusions based on the pilot and field test data.

The concern with "natural spoken language" is basic to assessment of listening and speaking. It would appear that there are two main types of spoken language--spontaneous and prepared. Spontaneous spoken language is generated as the speaker talks and is organized as the speaker's thoughts take shape. Wilkinson et. al. (1974) mention two kinds of spontaneous spoken language: "old organized speech"--the fluent speech used in ordinary conversation in which the probability of certain words occurring together is high, and "new now organizing speech"--the more hesitant speech used when speakers are thinking and choosing words as they go along. This speech is used to deal with cognitive problems requiring accurate solutions and therefore contains more redundancies, pauses, and false starts. On the other hand, prepared spoken language is organized prior to its use in a communication situation (as in giving a report from notes) and is generally in accord with a predetermined purpose and audience. Both spontaneous and prepared spoken language may reflect casual, informal, and formal levels of language use.

However, both these forms of spoken language are different from written language in several ways. In written language, sentences are carefully structured and linked. Written language is tightly organized and contains more objective details and fewer ambiguous ones than does spoken language. Spoken language is often loosely organized and may be characterized by incomplete or ungrammatical utterances and repetitive false starts. Written language contains punctuation to indicate vocal factors such as pitch, stress, juncture, etc. Thus an oral reader must follow the writer's vocal clues, whether or not he/she thinks these are appropriate or feels that the listeners will understand them. Vocabulary and writing conventions must also be used to provide information given by the speaker through vocal and nonverbal factors.

It would appear, then, that emphasis in oral communication assessment should be placed on both spontaneous and prepared oral language. However, since students must often deal with written material read aloud, it is necessary to assess competence in oral interpretation and listening to written language as well.

The results of Loban's longitudinal study of children's language are also of value in that they provide recommendations for assessing oral communication. (The results were reported in two monographs, written at various phases of the study. One was published in 1963 and the other in 1976.)

Loban's findings were reported in four categories: (a) fluency--readiness and smoothness of speech; (b) effectiveness and control of language; (c) functions accomplished by the use of language; and (d) interrelationships among the various aspects of language. Among the findings are the following: (a) power over language was found to increase through the successive development of control over meaningful forms--ability to handle pronouns, grammatical subordination, etc.; (b) no basic ability differences appeared among various ethnic groups; (c) subjects more proficient in language showed more sensitivity to its conventions and more frequently used language to express tentativeness; and (d) there was a positive relationship among the skills of listening, speaking, reading, and writing, suggesting that competencies in these areas reinforce each other. Among Loban's conclusions are these: (a) to achieve language flexibility pupils must apply whatever is studied to situations in which they have something to say, a deep desire to say it, and someone to whom they genuinely want to say it; (b) situation plays a large part in language fluency; (c) linguistic growth stages parallel stages of physical growth; and (d) more research on language development and use is needed.

Based on Loban's conclusions and recommendations (1976), it appears that oral communication assessments must take into account motivation and situation, as well as the physical, intellectual, and linguistic ages and stages of students. This is consistent with the recommendations of other researchers in the field of oral communication.

In summary, research reviews suggest the importance of the following considerations in assessing oral reception and expression:

1. Objectives for listening/speaking should be specified in terms of knowledges, skills, and attitudes; audience, situation, and purpose; cognition and affect; casual, informal, and formal language use.

2. Objectives should be concerned with spontaneous spoken language, prepared spoken language, and written language read aloud.
3. Assessment devices should be concerned with knowledges, skills, and attitudes; stimuli should be short, varied, and concerned with different audiences, situation, and purposes; casual, informal, and formal language should be reflected in these stimuli.
4. Assessment devices should provide for the evaluation of expression and reception of spontaneous spoken language, prepared spoken language, and written language read aloud.
5. Assessment devices should provide for the evaluation of a variety of oral language factors, including vocal, nonverbal, and verbal ones (vocabulary, syntax, etc.).
6. Assessment of oral communication competency should be conducted under controlled but natural testing conditions.
7. Attention should be given to examining available oral language test instruments to determine whether or not they meet the above criteria.

DESIGN OF THE STUDY

Rationale and Preliminary Planning

The products required by the contract between the Research Team and MACOSA were tests for use in assessing the achievement of objectives in speaking and listening in Grades 3, 6, 9 and 12. The following guidelines were given to the Research Team:

1. The tests will be appropriate for each of Grades 3, 6, 9, and 12 for each of speaking and listening.
2. The tests will be approximately one hour long for each target grade for each of listening and speaking. (This was later changed, with the approval of the Steering Committee, to "each target student" rather than "each target grade.")
3. The tests will be rated as "good" using criteria

> established by the Center for the Study of Evaluation (CSE), including criteria such as reliability, replicability, range of coverage, and score gradation where these apply.

These guidelines were affected by a number of factors relevant to the study design.

Prior to the inception of this study, MACOSA had contracted with the Edmonton Public School Board (EPSB) for the development of a set of listening/speaking objectives for elementary and secondary schools in Alberta. The objectives developed by EPSB reflected the broad scope of the language arts curriculum, with special emphasis on listening/speaking. Thus the Research Team felt that it would be essential as a first step in the study design to analyze the objectives as presented by the Steering Committee (these were selected from the EPSB objectives, and were prioritized and validated by the Steering Committee and Language Arts Policy Committee of the Department of Education); to categorize the objectives in terms of attitudes, knowledges, and skills; to indicate their relationships within these categories to spontaneous spoken language, prepared spoken language, and written language read orally; and to extend these categories and relationships in order to account for varied audiences, situational contexts, and listener/speaker purposes. Emphasis on written language (reading/writing) in Alberta schools had, to this point, rather effectively precluded the assignment of much instructional time to oral communication at these grade levels. Further, as is common throughout North America where oral communication is considered a strong part of the curriculum by individual teachers, schools, or systems, the current Alberta curriculum at all grade levels was more concerned with transactional and literary elements (report presentations, oral interpretation, etc.) than with spontaneous talk (conversation, discussion, brainstorming, etc.). Thus, the Research Team found themselves in the unusual position of designing tests to assess future knowledges and expectations in listening/speaking as well as current ones, and in the even more unusual position of assessing knowledge and skill development in areas with which teachers might well be lacking in knowledge and skill themselves.

Time and financial constraints had a strong effect on the development of the pilot, field, and final test instruments. Although the contract with the Research Team was not finalized until February, 1978, it was necessary that the entire project be completed by the end of September, 1978, in order to meet the requirements of MACOSA's mandate. Thus, a period of only three months was available to develop and administer test instruments before the close of the 1977-78 school year.

It had been originally planned by the Research Team that they would administer the pilot tests and that classroom teachers would administer the field tests. However, when the guidelines were changed to one hour of testing time per student, it was possible for the Research Team to do what they had wanted to do from the inception of the study--test speaking under individual rather than intact class conditions. In addition, the Research Team felt strongly that each speaker was entitled to an audience, and that without an audience the communication situation would be somewhat less than natural. Thus it was decided to administer the oral speaking skill portion of the test battery in small groups, using controlled pictorial stimuli at the Grade 3 and 6 levels and a specific topic at the Grade 9 and 12 levels.

However, the administrative plan for individual oral testing required a considerable amount of time on the part of classroom teachers. Further, it required space allotment outside of the individual classrooms and substitute teachers to cover classes while the classroom teacher was out testing the small groups. In addition, the testing procedure required that the talks made by students be taped so that they could be scored under both live and taped conditions. This was necessary because it seemed highly improbable that teachers could rank the necessary facets of talk during a one- or three-minute live presentation. Thus the scoring time required by teachers was lengthy, and the Steering Committee was concerned (quite justifiably) with teacher reaction to this time expenditure, especially toward the end of the school term. To add to this, the Research Team was concerned with the possible effect of bias, since teachers knew their students well and might rank them on what they felt the students could do rather than on what they did do during the testing situation.

It was therefore decided to involve a group of certified substitute teachers as examiners. These teachers administered all of the Pilot Tests and the Oral sections of the Speaking Field Tests, while classroom teachers administered the Listening Field Tests, the Written sections of the Speaking Field Tests and the Attitude tests.

The time, financial, and school situation constraints made it necessary to limit the range of communication events that could be tested. One area considered of vital importance, based on the review of the literature as well as the expertise of members of the Research Team, was that of small-group or interpersonal communication. It was agreed to eliminate this area from testing, not only because of the time factor but also because measures of such communication tend to be sophisticated and their interpretation therefore requires skills which the teachers could not be presumed to possess. Further, communicative interchange in small groups, both qualitative and quantitative, is largely dependent on classroom climate, and tests of such interchange might not measure a student's speaking ability or self-motivation, but rather a teacher's ability to establish a classroom climate conducive to free communicative interchange.

In the area of listening, it was clear that the most effective way to develop tape recordings for testing was to use professional actors and tape under professional studio conditions. However, both time and financial constraints, as well as the knowledge that the tapes would have to be made several times in order to incorporate revisions discovered to be necessary during the pilot and field testing, precluded the use of professional studio time during the making of the pilot and field test tapes. Thus sound reproduction during the pilot and field testing was not always the best quality. However test administrators did not consider this a problem. Furthermore, these conditions were remedied in the production of the final test instruments.

The guidelines presented to the Research Team called for consideration of the use of all or part of available and appropriate listening and speaking tests. In the area of speaking, a review of available test instruments indicated that none could comprehensively test the objectives developed by the Research Team and the Steering Committee. Thus all items, stimulus materials, and scoring devices were developed by the Research Team. (In some cases stimulus materials were published excerpts from stories, poems, or plays.)

Since attitudes toward communication were considered to be important, tests that might be related to the attitude objectives were examined. For this purpose, the Research Team finally selected the Personal Report of Communication Fear (PRCF) (McCrosky, 1977) for administration to Grades 3 and 6. The Personal Report of Communication Apprehension (PRCA), by the same author, was used for administration to Grades 9 and 12.

Design Steps

In order to expedite the study the Research Team divided its work into 20 Design Steps. These are briefly presented below.

1. Selection and revision of relevant, high-priority measurable objectives.
2. Review of existing listening/speaking tests and related literature.
3. Structuring of the listening/speaking domain. (See Figure 1.) A detailed analysis of this Figure is presented in the Final Report.
4. Devising of indicators of acceptable performance.
5. Organization of selected objectives by attitudes, knowledges, and skills.
6. Evaluation of existing tests.
7. Construction of the Pilot Test item pool.
8. Construction of the Pilot Test Blueprints.
9. Construction of the Pilot Test instruments.
10. Administration of the Pilot Test instruments.
11. Analysis of the Pilot Test data.
12. Revision of items based on Pilot Test data.
13. Development of the Field Test Blueprints.
14. Construction of the Field Test instruments.
15. Development of the Field Test Sampling Plan.
16. Administration of the Field Test.
17. Development of the Field Test Data Analysis Plan.
18. Preparation of the Final Test Instruments.
19. Preparation of the Reporting Plan, development of a draft of the Technical Report and Condensed Report for submission to the Steering Committee for their consideration and approval.
20. Development of the Technical Report and Condensed Report (Conclusions and recommendations are included in Chapter 5 of the Technical Report)

Thus Steps 1-11 were concerned with identifying and organizing the listening/speaking objectives, reviewing the literature, constructing the Pilot Test items, constructing the Pilot Test instruments, administering these instruments, and analyzing the data obtained from the test administration. Steps 12-20 were concerned

THE COMMUNICATION PROCESS

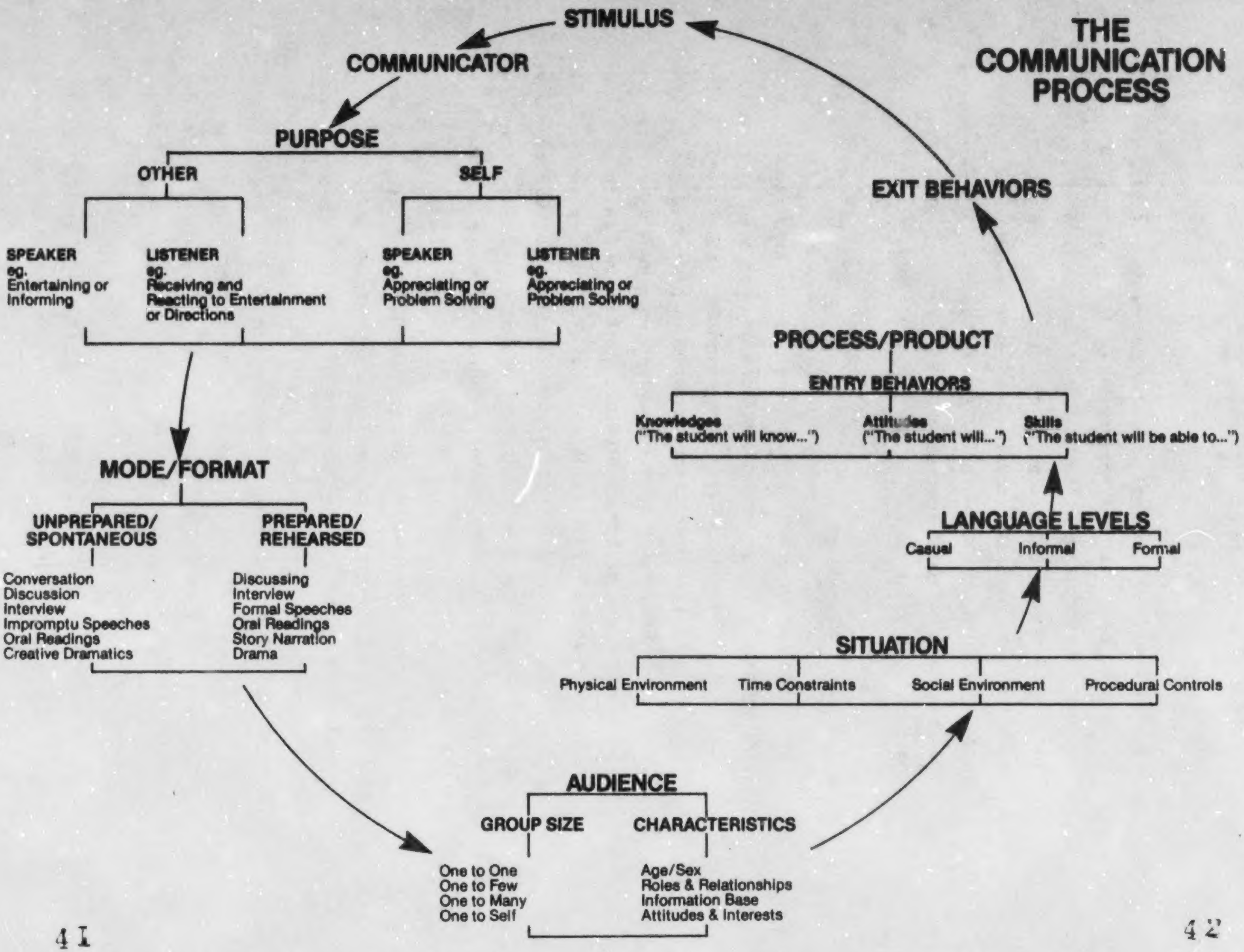

STIMULUS
COMMUNICATOR
PURPOSE
OTHER
SELF
SPEAKER eg. Entertaining or Informing
LISTENER eg. Receiving and Reacting to Entertainment or Directions
SPEAKER eg. Appreciating or Problem Solving
LISTENER eg. Appreciating or Problem Solving
MODE/FORMAT
UNPREPARED/ SPONTANEOUS
Conversation
Discussion
Interview
Impromptu Speeches
Oral Readings
Creative Dramatics
PREPARED/ REHEARSED
Discussing
Interview
Formal Speeches
Oral Readings
Story Narration
Drama
AUDIENCE
GROUP SIZE
One to One
One to Few
One to Many
One to Self
CHARACTERISTICS
Age/Sex
Roles & Relationships
Information Base
Attitudes & Interests
SITUATION
Physical Environment
Time Constraints
Social Environment
Procedural Controls
LANGUAGE LEVELS
Casual
Informal
Formal
PROCESS/PRODUCT
ENTRY BEHAVIORS
Knowledges ("The student will know...")
Attitudes ("The student will...")
Skills ("The student will be able to...")
EXIT BEHAVIORS

Figure 1

with revising the Pilot Test stimulus materials and items where indicated by data analyses, constructing additional items where necessary, constructing Field Test instruments, administering Field Test instruments, analyzing the data obtained from the test administration, revising stimulus and pilot items where indicated by data analyses, constructing additional items where necessary, constructing final test instruments, and preparing the Technical Report and Condensed Report.

Selection of Objectives

A first set of prioritized objectives was presented by the Research Team to the Steering Committee and approved by them in March, 1978. These objectives were classified by attitudes, knowledges, and skills. Justification for this procedure can be found in the literature, where oral communication generally includes both cognitive and affective elements, and where cognitive elements usually contain both knowledges and skills. Further justification can be found in the fact that courses in speech, at whatever level, contain both affective and cognitive elements, with specific concern directed to the development of both knowledges and applicational skills.

It was also hypothesized by the Research Team that a student might know the principles of message preparation and reception or of oral speaking and yet not be able to apply them in the speaker-audience situation. If this were the case, it might be due to such aspects of lack of knowledge of factors involved in oral communication, lack as experience in oral performance, or fear in the speaking situation. Thus it was essential to isolate these various factors through evaluation of knowledges, skills, and attitudes in oral communication.

In a document entitled Specific and Prioritized Objectives for Listening and Speaking, prepared for the Steering Committee by the Research Team and presented to them on March 7, 1978, the sets of categories used to classify the knowledges, skills, and attitudes dimensions were enumerated and described.

Once the objectives had been selected and prioritized, it was necessary to group them into categories or domains. This was necessary to facilitate data analysis and interpretation. Both the objective lists and domain structures were revised several times. The final domains for speaking and listening are presented below.

Listening	Speaking
1. Environmental Sound Factors	1. Voice Production Factors
2. Voice Production Factors	2. Nonverbal Factors
3. Linguistic Factors: Words	3. Linguistic Factors: Speech Sounds
4. Linguistic Factors: Sentences	4. Linguistic Factors: Words
5. Organizational Factors: Plan Analysis	5. Linguistic Factors: Sentences
6. Organizational Factors: Literal Comprehension	6. Organizational Factors: Ideational Fluency
7. Organizational Factors: Interpretive Comprehension	7. Organizational Factors: Plan Analysis
8. Organizational Factors: Critical Comprehension	8. Organizational Factors: Organization of Details

Item Development

The flow chart provided in Figure 1 indicates how specific objectives can be related to any stimulus, audience, mode, purpose and communicator, at any level of stimulus and communicator sophistication. This flow chart served as the basis for the development of the stimulus materials and items. A number of factors had to be taken into account in item preparation.

1. There was a need to prepare more items than necessary for the Pilot Tests so that a

sufficient number of items would be available for the field test.

2. Where possible there were to be at least two items for each objective.
3. The test content was to reflect three types of language: written, spontaneous, and spontaneous planned. An appropriate balance among these types of language as reflected in the Test Blueprints was agreed upon.
4. Stimulus materials should be based on the kinds of activities students might naturally encounter in a classroom setting.
5. For purposes of analysis, and to standardize testing procedures, it was decided that a multiple-choice format, using three options in each item, would be used in all tests. Rating scales in speaking tests were originally constructed using a 1-0 format--that is, the rater merely indicated the presence or absence of a quality or ability. (For the field testing this was changed to a scale on which raters could score on a 1 to 5 basis. Again this made a more incisive analysis possible.)
6. The review of the literature indicated that published tests, which were all norm-referenced, tended to use a large number of items to measure only a few objectives. Thus the Research Team designed all stimulus materials and test items used in this study.

Using the flow chart in Figure 1, the prioritized objectives, and the domain structures referred to above, the Research Team constructed stimulus materials and items for the Pilot Tests. In addition, they prepared Test Blueprints and Objective/Items Matrices for each grade level for both listening and speaking. Pilot testing was then conducted; statistical analyses were made; and the tests were revised for further administration. The pilot and field testing procedures, the analysis of data, and the interpretation of the results are described below.

PILOT TEST ADMINISTRATION AND DATA ANALYSIS

The purposes for pilot testing may be summarized as follows:

1. To determine if the subjects can understand the directions;
2. to determine if the administrative procedures are adequate;
3. to determine if the options in multiple choice items are working properly;
4. to determine the amount of time required for each item or group of items;
5. to determine if items discriminated or were too difficult or too easy (floor effects and ceiling effects);
6. to determine the practicality and reliability of the scoring system;
7. to provide baseline data for the preparation of the field test instruments.

Sample

The Pilot Test sample was the same for both the Listening and the Speaking Test administration. At the Grade 3 level two intact classes of pupils from Calgary and two intact classes from Okotoks were used (N = 117). The same schools were used for Grade 6 (N = 124). The Grade 9 sample (N - 101) consisted of two classes from Calgary and two from Airdrie. In Grade 12 students from Calgary and Springbank (N = 82) were tested.

Procedures

After pilot schools had been selected, all principals and teachers were contacted; raters were hired and trained; and materials were distributed to schools. The pilot testing was conducted in May, 1978. It should be noted that all test administrators were certificated teachers who were currently on the Calgary

Board of Education Substitute Teachers List. All tests were administered by these teachers who were hired and trained by the Research Team.

Analysis of Pilot Test Data

As indicated earlier there were a number of purposes associated with the administration of the Pilot Tests. Thus statistical analyses were performed to determine: a) if the options in the multiple choice items were working properly; b) if the items were too difficult or too easy; c) if the items discriminated among students; and d) if the ratings and tests themselves were reliable. In order to see to what extent the Pilot Tests met these criteria, items analyses, choice distributions, and inter-rater reliability coefficients were calculated for each of the tests. The results of these analyses are summarized below.

1. Items were evaluated for difficulty level. A cut-off of .30 to .80 in the Difficulty Index was arbitrarily established as a criterion. That is, items which were correctly answered by fewer than 30 per cent or more than 80 per cent of the students would be subject to revision or deletion.
2. Next, items were evaluated on the basis of their usefulness in discriminating between strong and weak students. A Discrimination Index level of .20 was applied.
3. Items in which options attracted fewer than five per cent of student responses were also subject to revision.

Using this information as well as comments made by teachers, administrators, etc., the Research Team were able to revise the instruments for field testing. We discovered that, indeed, there were a number of items that did not meet established criteria. Since the Pilot Tests were longer than required, however, it was possible to delete a number of weak items and still meet the requirements of a balanced test as well as of including at least two items to measure most of the objectives. Some items were retained even though they were easy because they measured high-priority objectives. The importance of objectives thus remained as the prime criterion in the item selection process for this test.

Rater Reliability

In order to test the reliability of the ratings made by the judges on the Speaking production items, a sample of the taped oral readings and speeches was selected and rated. This was done by selecting every fifth tape in each grade. Two raters--substitute teachers who had administered the tests but who had not rated the students whose tapes were selected--rated each tape. It should be noted that items which required visual contact between rater and student were not included in this analysis. The results of these ratings were correlated to arrive at reliability figures. Correlation coefficients ranged from .75 to .83. The average correlation was .78 for all four grades.

In order to determine whether there was adequate validity in the ratings, we also looked at choice distributions. There appeared to be considerable spread in the ratings, indicating that raters did not simply assign the same score to students within any one scale. We did note, however, that there were several items which raters frequently omitted. These items were revised by providing more adequate rating criteria and were included again in the Field Tests because it was felt that they measured important objectives.

We also considered the feasibility of changing the form of the Rating Scales. In the Pilot Test a 1-0 (Yes, No) rating scheme was used. We felt that teac[illegible] would be unable, and perhaps unwilling, to use a five-point scale for the number of items that were required.

The results of our analysis and comments from the teachers suggested that a five-point scale would be feasible, and so the Rating Scales were adjusted to this format for the Field Tests. At the same time we realized that this procedure might possibly reduce the reliability of scoring.

FIELD TEST ADMINISTRATION AND DATA ANALYSIS

Sample

The Field Test sample was the same for both the Listening and Speaking Test Administration. All students were from Red Deer and Didsbury where intact classes were used. Thus 210 students in Grade 3, 235 students in Grade 6, 251 students in Grade 9, and 168 students in Grade 12 were tested.

Procedure

The procedures for field testing were substantially the same as those used in pilot testing. Contacts were made with the schools; the nature of the testing program was discussed with teachers and administrators; the Listening, Written Speaking, and Attitude tests were administered by classroom teachers; the Oral Speaking Tests (Rating Scales) were administered by substitute teachers hired and trained by the Research Team; tests were collected and coded for key punching; and the data were submitted to the computer for analysis.

Analysis of Field Test Data

The Technical Report presented the results of the data analysis in detail. Only a brief summary of each analysis, together with the conclusions which may be drawn from them, will be presented here.

Item Analysis. As in the pilot testing phase we were concerned with item difficulty and discrimination. Using a cut-off of .80 for difficulty and of .20 for acceptable discrimination, we found that most of the items were acceptable for use in the Field Tests. In some cases items were changed by making options more attractive. In some of the Rating Scales additional criteria were provided for more clarity in the items.

An analysis of difficulty and discrimination was also carried out on the basis of male and female samples. No significant differences were found, although there was a tendency for females to achieve higher scores than males.

The conclusion from these analyses was that with minor revisions these tests were now ready for administration to larger samples of Alberta students and that they were sufficiently reliable and valid to meet the requirements of a "good" test. While it might be argued that there were still some easy items, especially in the Written Speaking Test, it must be remembered that the primary concern of a criterion-referenced test is its ability to provide information about levels of achievement in specific areas rather than to produce some overall score. Thus easy items were retained if it could be demonstrated that important, high-priority objectives were being measured.

Descriptive Statistics. Descriptive statistics were presented to indicate means, standard deviations, and reliabilities for each Listening and Speaking subtest and for the total tests. A satisfactory spread of scores was obtained; the total tests appeared to be at about the right level of difficulty; and reliabilities--ranging from .43 to .67 in Listening and from .39 to .64 on the Written Speaking test--appeared acceptable for a criterion-referenced test. (It should be noted that these reliability coefficients are low when compared to those reported for standardized tests. Such tests are, however, composed of homogeneous items measuring only a few objectives. A more persuasive reliability estimate would be obtained by giving the same test to the same group with a time interval between the two administrations.)

Again, on the basis of these data, the Research Team concluded that the tests were ready for administration to larger groups of students.

Correlational Analyses. Several types of correlational analyses were carried out to determine the extent to which subtests related to each other and the extent to which listening, speaking and attitude were related as a whole. Of most interest were the correlations which appeared when we correlated the item subsets which had been classified into domains. (For example, one domain would be Voice Production Factors. Another would be Organization of Details.) The results of these analyses are summarized below.

1. The domain sub-scales in listening correlated at a very low level with each other--that is, each subscale was relatively independent from the others. This was true at all grade levels in listening and supported the notion that we had indeed produced criterion-referenced tests rather than ones which measured general listening ability.
2. The domain sub-scales in the Speaking Oral Tests were more highly correlated with each other. We concluded that this was largely due to the fact that the same raters rated each student and that such ratings might be caused by a "halo" effect. The fact that the item analysis data did reveal a considerable spread of scores in the ratings did suggest, however, that they were acting as acceptable measures.
3. The Speaking Written Tests tended to correlate more highly with the Listening Tests. This suggested that a different ability, more like that required in listening, was operating here.
4. There was a low negative correlation between Listening Total and Speaking Total. While this correlation was low; it was consistent across all grades. The Research Team tentatively concluded that good listeners may be poor speakers.
5. There was no correlation between the attitude measure used and any of the other tests. While

> we have no evidence to suggest that the scales we used to measure attitude do not, in fact, perform this function, we did conclude that the affects of such anxiety cannot be traced by the methods we used, nor does such anxiety appear to related to the types of activity involved in the tests.

In summary, these analyses indicate that the tests which were constructed in this study are valid in terms of measuring a variety of objectives; no reasonable way of re-designing the tests to make them shorter is apparent; and there appears to be no reason that the attitude measure used here should be given again.

Factor Analysis

Since the domains, as noted above, were arrived at by priori process, the Research Team felt that it would be useful to perform factor analyses on the Listening, Speaking, and Attitude tests by grade. Test statistics, item analysis data, and correlation data suggested that a male-female distinction in such an analyses would not be productive. The limited sample size, too, might lead to spurious results. Thus, a first-run factor structure based on the total population at each grade level was felt to be adequate and defensible. The results reported below are based on the use of the Varimax Rotation using an SPSS program.

Analysis of the Grade 3 Listening Test yielded six factors which accounted for 76 per cent of the variance. These factors were labelled and compared to the original domain structure. A reasonably close match was obtained, and it was concluded that the domains as originally conceived formed logical groupings of items for test interpretation.

Attempts were made to perform similar analyses of Grade 6, 9, and 12 Listening Test data. No reasonable factor structure could be found. Thus a large number of factors was required to account for substantial amounts of variance. We concluded that the items on the test were heterogeneous--that is, they required a large number of abilities and skills which could not be grouped into meaningful categories. Since this was the case, no comparison of factor groupings to original domain structures could be made.

In the Speaking Tests, clear-cut factor structures emerged in Grades 3, 6, and 9. In each case six factors accounted for at least 90 per cent of the variance. In Grade 12 a number of items, specifically those on the Written Test, did not load on any of the primary factors. This latter finding suggested that at this level the skills required to perform well on this test are of a heterogeneous nature and are largely dependent on the background and ability of individual students. The match between the obtained factors and the original domain structures was good. Again, we concluded that the groupings of items in the original domains was a satisfactory base for interpretation.

CONCLUSIONS AND RECOMMENDATIONS

On the basis of the data analysis the Research Team concluded that, with minor revisions, the field-tested instruments could now be used as Final Tests for large-group testing in Alberta schools. It should be noted, however, that these Final Tests, especially the Speaking Tests, are relatively novel, as was the developmental approach generally taken by the Research Team. For this reason, we make a number of recommendations about further test development. The next section summarizes our conclusions and provides suggestions for the use of the tests as well as for further test development.

Future Test Administration

Conclusion 1: Item analysis and other data indicate that the tests in both listening and speaking appear to be valid measures of the prioritized objectives selected for this study. It should be noted, however, that testing in these areas is a relatively new development. In fact, to our knowledge, a criterion-referenced test in these areas does not exist. Because of this, our conclusions and recommendations may appear to be somewhat more tenuous than would be those of test makers in the more developed areas. For this reason, and because the tests were intended as Field Tests only and therefore were used on a limited sample, a number of caveats are stated below:

A. The objectives used as a basis for item design were those selected and prioritized by the Research Team and the Steering Committee. (See the Preface and Chapter 2 of the Technical Report for a description of the objectives development process.) Thus they do not represent the total scope of objectives that one might wish to include and assess in the listening and speaking areas of the language arts/English programs in elementary and secondary schools.

B. In the Oral Speaking Test, some of the sub-scales appear to account for substantial portions of the variance. Specifically, Voice Production, Linguistic Factors (Words) and Organization of Details correlate highly with Speaking Total. (It is possible to suggest that some sub-scales could be dropped. This would have the effect, however, of omitting items which measure high-priority objectives, and we do not feel that it is wise to make any deletion of items at this time.) It may also be said that the Written Speaking Test contributes little to the variance. It is our feeling, however, that students need to perform the Written Speaking Test in order to prepare adequately for actual production. Furthermore, this portion of the battery measures a distinct ability which relates clearly to at least some high-priority objectives not measured elsewhere.

C. Estimates of reliability were based on inter-item data analysis only. This, in effect, is a measure of homogeneity of items. No data were available which were based on parallel forms which could provide test-retest reliability estimates.

D. Rater reliability in the Oral Speaking Test was based on Pilot Test data. While there was some difficulty in establishing an adequate level of rater reliability, this had more to do with the quality of the audio tapes used than with the teachers' rating abilities. At times the quality of these tapes interferred with effective rating. An alternative might have been to have several raters in the room where the students were responding. However, this would probably have created a very artificial condition and would have interferred with test results. (Rooms with one-way window facilities might solve this problem.)

E. One of the areas with which we were concerned and which proved to be extremely difficult was that of devising items which would show developmental trends from grade to grade. In fact, a number of items were designed which "crossed" grade levels. They appeared to have face and content validity, but statistical analysis did not yield sufficient information to indicate clearly that they gave valid evidence of developmental trends.

Recommendation 1-1

We recommend that the Speaking and Listening Final Tests be given to larger samples of the Alberta school population, and that further analyses be performed on the results with a view of establishing acceptable performance criteria.

Recommendation 1-2

Alberta Education should determine appropriate techniques for ensuring that data derived from future testing programs in listening and speaking are carefully analyzed, collated, and "banked" so that useful comparisons can be made over time. It would also be useful if further testing in all language arts areas were performed on the same sample. This would make further correlational analysis possible.

Recommendation 1-3

In an effort to establish developmental trends, it may be useful to administer each Listening and Speaking Test at one grade level above and below the level for which the test is designed. Thus the Grade 3 test would be given to Grades 2 and 4, the Grade 6 test to Grades 5 and 7, etc. This would, of course, be possible only in the downward direction for the Grade 12 tests. Such test administration would also prove useful in the development of performance criteria at all grade levels.

Recommendation 1-4

No possibility for shortening the Listening or Speaking Tests appears appropriate. Since these tests are based on stimulus subscales, and since these sub-scales contain a mixture of items/objectives measures, they must be given in intact form. Thus we recommend that the entire Listening Test and the entire Speaking Test for each grade level be given as presented. (However, each test in its entirety need not be administered in one sitting.) Additionally, further study is required to determine if preparation for oral production makes a real difference in the way students actually perform.

Future Testing and Data Analysis

Conclusion 2: Statistical analysis of data largely revolves around judgments made by the researchers. Many reasonable arguments could be made for using techniques other than those reported in this study. For example, we would draw attention to the correlational analysis in which sub-scale correlations were based on data derived from item measures structured on the basis of a-priori domains. If these domains were not originally correctly conceived, this analysis might not be adequate. The factor analytic technique may also be questioned. We felt that the use of the Pearson Product-Moment Correlation Coefficients was acceptable, given the nature of the data. Tetrachoric correlations might, however, be more acceptable. Finally, basing discrimination indices on stimulus sub-scales may also be questionable. However, this appeared appropriate for the purpose of field testing, particularly in view of the fact that criterion-referenced tests were being constructed to which traditional analyses did not necessarily apply.

Recommendation 2-1

Further test reliability data should be collected. Since it is extremely difficult to construct parallel forms, a sample of students at each grade level should be selected and the tests should be given at two different points in time so that test-retest reliability can be established.

Recommendation 2-2

More data are needed to establish rater reliability. We recommend that this be dealt with as a separate study in which raters observe students from some area where they cannot be seen by the students taking the tests.

Recommendation 2-3

Continuing efforts should be made to devise items which will reveal developmental trends in the areas of listening and speaking. It is quite likely that this may need to involve the development of additional and perhaps even more specifically stated objectives.

Recommendations 2-4

Further statistical analyses should be done when data are derived from the Final Test samples. It is also suggested that analyses involving correlational techniques be used in order to devise sub-scales other than those based on a-priori domains.

Recommendation 2-5

Further factor analytic analyses should be undertaken based on the larger Final Test (and/or other) samples and should use other forms of correlation--for example, tetrachorics. Such samples would also permit the study of special sub-groups--that is, male/female; rural/urban; high/low abilities; etc.

Recommendation 2-6

When larger samples have been tested, it should be possible to perform analyses more suited to criterion-referenced testing. While these methods of analysis (item curve characteristics, etc.) are in the neonate stage, some are becoming sufficiently well explicated to be of use. At the same time, it is our feeling that traditional analysis techniques will still be useful in providing full and complete data upon which to base appropriate conclusions.

Future Test Development

Conclusions 3: Since the tests developed in this study appear to be viable, it becomes necessary to look at possible further test development. The tests constructed for this study are limited necessarily to the measurement of those objectives which formed the measurement of those objectives which formed the basis of their design. Thus they cannot reflect all of the current listening and speaking aspects of the English/language arts curriculum, nor do they purport to reflect what might reasonably be included in such a curriculum. Nevertheless, the items were based on objectives which received high priority in the literature as well as on those which were perceived by competent judges and experts to be valid and important.

Recommendation 3-1

While objectives for this study were carefully selected, further energy should be expended in continuous review and upgrading of objectives. A mechanism is needed for ensuring that such objectives become a part of the language arts/English curriculum. Such ongoing development would also require constant review and development of measurement techniques.

Recommendation 3-2

The Research Team feels that the *Listening Comprehension Tests* developed in England by Wilkinson, Stratta, and Dudley (1974) would be a most useful addition to a Listening Test battery for Alberta schools. While these tests measure objectives not specifically dealt with in this study for reasons described in the *Technical Report*, they are based on objectives considered highly relevant and valid by experts in the language arts field. If this test is to be used, Canadianization is essential.

Recommendation 3-3

The tests provided in this study are useful as initial tools for identifying students who have difficulties in listening and speaking. However, the tests were not primarily designed as diagnostic tests of listening/speaking difficulties, and therefore other tests suitable for such purposes need to be developed and/or selected, and made available to teachers.

Attitude Testing

Conclusion 4: The data in this study indicate that there is no relationship between the attitude measure and any of the other measures used in this study. (This was a rather surprising finding because it seemed reasonable to expect that a measure of communication anxiety should correlate with the area of speaking, at least.) A number of reasons for this lack of relationship have been suggested in the Technical Report. Since the measure of attitude used was a very "pure" one--that is, it loaded on only one factor--it may be that a correlation of any magnitude could not reasonably have been expected. In effect, then, we can conclude that the Attitude Tests used may very well measure communication anxiety, but that such anxiety does not necessarily affect a person's ability to perform the tasks required in the Speaking and Listening Tests as structured in this study.

Recommendation 4-1

We do not recommend the use of the Personal Report of Communication Fear or the Personal Report of Communication Anxiety (McCrosky, 1977) as part of the testing program in listening and speaking. These tests may well be useful as measures of communication anxiety, but their use should be limited to measuring only this aspect of student development.

Recommendation 4-2

If MACOSA deems it necessary to provide attitude tests for the listening/speaking areas of the language arts curriculum, a specific program for developing such tests should be undertaken. It is our opinion that tests of attitudes toward school, such as the ones currently being developed for MACOSA, might well be feasible.

Reporting Test Results

Conclusion 5: One of the concerns in any province-wide testing program relates to dissemination and security. In our opinion, it would not be wise to release any of the Pilot, Field, or Final Listening or Speaking Test items or stimulus materials. These are, after all, intended for assessment rather than instructional purposes.

Recommendation 5-1

No tests or stimulus materials should be released, and rigid control should be exercised over copyright, test administration, return of copies of the tests tapes, etc.

Recommendation 5-2

The data collected as a result of the administration of the Final Tests should not be disseminated to participating schools or other interested parties until acceptable performance criteria have been established. Release of such data should be carefully controlled, and clear and precise explanatory material should be provided indicating the purposes, functions, and limitations of the test instruments.

Recommendation 5-3

The Research Team recommends that the tests submitted with the Technical Report be administered to a large sample of Alberta students as indicated in the recommendations outlined above. As a result of such a testing program, performance criteria could be established and the tests could be used as a criterion measure whenever Alberta Education or a school system requests such an evaluation.

Recommendation 5-4

Alberta Education should prepare a statement of those objectives it thinks are appropriate as part of the school program in listening and speaking at the elementary and secondary grade levels. Such a list should indicate which objectives are measured by the Final Tests prepared for this study. This statement should be made available to school systems throughout Alberta for use in future curriculum planning.

Teacher Training

Conclusion 6: Since the teaching of listening and speaking has not been emphasized until recently in the Alberta school language arts/English curriculum, and since it has not been a major element in teacher education programs, it is important to ensure that teachers are provided with the requisite skills necessary for effective initiation and development of the listening and speaking components of the school program.

Recommendation 6-1

Pre-service and in-service training at all levels should be made available to teachers concerned with providing language arts/ English instruction in elementary and secondary schools, in order that they can develop and expand the requisite skills for teaching and diagnosing listening and speaking knowledges and skills as well as for integrating listening and speaking into the total language arts/English curriculum.

Recommendation 6-2

Pre-service and in-service training in the area of functional communication skills should be made available to all teachers, regardless of their curriculum speciality.

Recommendation 6-3

When the Final Listening and Speaking Tests are in general use in the province of Alberta, pre-service and in-service training should be provided to teachers in both the administration of the tests and the interpretation of test results. Further, training should be provided to teachers in the use of the test results in curriculum planning at the individual, class, school, and system levels.

REFERENCES

Devine, T. "Listening: What Do We Know after Fifty Years of Research and Theorizing?" Journal of Reading, Vol. 21, January, 1978.

Edmonton Public School District. Listening and Speaking Objectives. Report to the Minister's Advisory Committee on Student Achievement (MACOSA). Alberta Education, 1979.

Landry, D. "The Neglect of Listening." Elementary English, Vol. 46, 1969, pp. 599-605.

Larson, C., P. Backlund, M. Redmond, and A. Barbour. Assessing Functional Communication. Falls Churc, Va.: Speech Communication Association, 1978.

Loban, W. The Language of Elementary School Children. Urbana, Illinois: National Council of Teachers of English, 1963.

__________. Language Development: Kindergarten Through Grade Twelve. Urbana, Illinois: National Council of Teachers of English, 1976.

Marchak, N. et. al. Assessing Communication Skills: A Review of the Literature. Report to the Minister's Advisory Committee on Student Achievement (MACOSA). Alberta Education, 1979.

Mead, N. Issues Related to Assessing Listening Ability. Unpublished paper, Toronto: American Educational Research Association, 1978.

Plattor, E. Assessing Listening Skills--An Examination of Selected Research Reports. Report to the Minister's Advisory Committee on Student Achievement (MACOSA). Alberta Education, 1977.

Wilkinson, A., L. Stratta, and P. Dudley. The Listening Comprehension Tests. Toronto: Macmillan of Canada, 1974.

__________. The Quality of Listening. Toronto: Macmillan of Canada, 1974.

TEST DEVELOPMENT FOR ASSESSING ACHIEVEMENT IN LISTENING AND SPEAKING

PART II

TECHNICAL REPORT

PREFACE

The study described in the following pages had as its major objective the development of test instruments designed to assess Alberta student achievement of selected objectives in listening and speaking at the Grade 3, 6, 9, and 12 levels.

The study was conducted for the Minister's Advisory Council on Student Achievement (MACOSA) by principal investigators E. Plattor, W. Unruh, and L. Muir during the months of February-October, 1978. K. D. Loose of the Faculty of Continuing Education, University of Calgary acted as statistical consultant. The study was administered by the Planning and Research Branch of Alberta Education, under the supervision of a MACOSA-appointed Steering Committee comprising B. Gommeringer (Chairperson, Alberta Education), W. Laing (University of Alberta), C. Rhodes (Alberta Education), and W. Washburn (Calgary Board of Education). H. Sherk and M. Thornton, representing Alberta Education, served as consultants to the Steering Committee.

In June, 1977, B. Gommeringer, Regional Office Consultant with the Curriculum Branch of Alberta Education, was requested to prepare a report for MACOSA outlining the ways, means, and costs of conducting a province-wide assessment of student achievement in listening and speaking. This report was developed by B. Gommeringer, E. Plattor, and P. McFetridge (University of Alberta). Specifically the report was designed to: (a) indicate the desirability or feasibility of assessing the achievement of Alberta students in listening and speaking; (b) provide an outline of the tasks required in conducting such an assessment; and (c) provide an estimate of the costs involved in conducting such an assessment. The report indicated that, in the opinion of its authors, it was indeed desirable and feasible to conduct such an assessment and outlined possible activities, techniques, and costs involved in its development. The report was submitted to MACOSA on July 13, 1977 and changes were made in the objectives based on the recommendations of this panel. Considerable attention was given to this project by the Curriculum Branch of Alberta Education. Upon completion of the task, the objectives were validated and prioritized by the Language Arts Coordinating Committee of the Curriculum Branch, and then by the Steering Committee. The objectives were further refined and modified by the Research Team, based on their review of the literature in listening and speaking as well as on discussions with the Steering Committee and other knowledgeable colleagues.

The Research Team is indebted to the Steering Committee for their encouragement and insightful recommendations. The advice and interest of colleagues at the University of Calgary, as well as colleagues in Alberta schools and Departments of Education are also gratefully acknowledged. Discussions with educators in other provinces and in the United States were also most helpful in the development of the objectives and the tests themselves. The cooperation and encouragement provided by the administrators, teachers, and students who participated in the study are also very much appreciated. Finally, a special debt of gratitude is owed to the Research Assistants, whose dedication to their task was often far above and beyond the call of duty.

Chapter 1

INTRODUCTION TO THE STUDY

The Minister's Advisory Committee on Student Achievement (MACOSA) was established by Ministerial Order in October, 1976, to identify solid evidence of the level of achievement in Alberta schools (Alberta Education, 1977). The Committee's mandate included: (1) studying the elements of student achievement, including definition, purpose, principles, and assumptions underlying achievement; (2) consideration of ways and means to assess student achievement; (3) consideration of methods for maintaining or improving achievement; and (4) recommendations on these matters as based on findings.

In order to fulfill this mandate, MACOSA commissioned a number of studies in various achievement categories. On February 7, 1978, Alberta Education contracted the University of Calgary, and specifically the principal investigators E. Plattor, W. Unruh, and L. Muir, professors in the Department of Curriculum and Instruction and of Educational Psychology, to design and/or adapt tests for assessing the achievement of selected objectives in listening and speaking for Grades 3, 6, 9, and 12.

The investigators worked under the supervision of a MACOSA-appointed Steering Committee to: (1) prepare lists of measurable listening/speaking objectives based on ones prioritized and validated by the Steering Committee in cooperation with the Curriculum Branch of Alberta Education and their Language Arts Coordinating Committee; (2) review related literature in listening/speaking; (3) construct an item pool for assessing selected listening/speaking objectives; (4) construct and administer pilot test instruments to a sample of students at Grades 3, 6, 9, and 12; (5) analyze the data obtained from the pilot tests and interpret these data in order to provide baseline information for the production of field test instruments; (6) construct and administer field test instruments to a sample of students at Grades 3, 6, 9, and 12; (7) analyze the data obtained from the field tests and interpret these data in order to provide baseline information for the production of final test instruments; (8) construct final test instruments in listening and speaking at the Grade 3, 6, 9, and 12 levels; (9) compile into a Technical Report all of the information obtained from the pilot and field testing together with recommendations.

In a letter of December 16, 1977, the principal investigators were invited to submit a Prospectus relative to revisions and/or development of tests to measure achievement of objectives in listening and speaking. In this letter, it was pointed out that MACOSA had already commissioned an initial study in the areas of listening and speaking. The products of that activity, which was at that time being carried out by the Edmonton Public School District, were to be validated statements of objectives for listening/speaking. This study was completed, and the statements of objectives were presented to the principal investigators when they began the development of the pilot test instruments.

A Prospectus entitled Assessing Achievement of Selected Objectives in Listening and Speaking was submitted by the principal investigators to Alberta Education on January 10, 1978. This was followed by a request from Alberta Education for a formal Proposal for the study. A Proposal, also entitled Assessing Achievement of Selected Objectives in Listening and Speaking, was submitted on February 1, 1978. This Proposal was accepted by Alberta Education, and a contract between Alberta Education and the University of Calgary for a project entitled MACOSA-Developing and Validating Achievement Tests in Speaking and Listening was finalized on February 7, 1978.

During the Steering Committee meeting of March 7, 1978, a number of guidelines were agreed upon in order to clarify expectations regarding test development. These were:

1. Length of the Tests. The original terms of reference for the study required that the maximum length of the tests at all grade levels was to be approximately one hour in both speaking and listening. At this meeting, it was agreed that the basis for establishing the 60 minutes would be the time involved in testing an individual student, rather than the time taken for testing a total class. It was further agreed that the test package could include a medley of tests or testing procedures.
2. Use of Audio and Visual Equipment. Wherever possible, it was agreed that technological aids would be considered in presenting testing tasks and recording observations. Thus, the listening tests were presented to students through the use of audio tapes prepared by the investigators. In addition, all of the talks given by students during the administrations of the speaking tests were audio taped for scoring purposes.
3. Testing in Natural Situations. Wherever possible, it was agreed that the test situations/observations would be structured to capture the performance of students in natural settings. Emphasis would be on classroom situations both formal and semi-formal, including small group discussions, impromptu reports, show-and-tell activities, etc.

During the Steering Committee meeting of April 6, further guidelines were agreed upon.

4. Statistical Analyses. Since the tests were developed as criterion-referenced instruments, objectives would be placed in domains for purposes of analyses. These analyses, however, would include traditional analyses as well as new iterative approaches to assessing the value of items and in processing results.

5. Commonality of Stimuli. Wherever possible and appropriate, common stimuli would be used among two or more of the grades so that comparison across grades could be attempted.

6. Speakers and Audiences. As a general rule, when stimulus items required the recording of speakers, such speakers were limited to teachers or students. Listeners were defined as classroom audiences.

7. Knowledge, Skill, and Attitude Testing. Wherever possible, knowledges and attitudes in listening and speaking would be tested with pencil-and-paper instruments. Skills would be tested directly, by speakers giving speeches, and listeners responding to actual talk provided by tape recordings.

PURPOSES OF THE LISTENING/SPEAKING ASSESSMENT

The test instruments were designed for the following purposes:

1. To provide for program evaluation through the collection of baseline date in listening and speaking by the province, school systems, and/or individual teachers. For this purpose, measures of group achievement would be the primary concern.

2. To provide for diagnostic assessment of individual students by their teachers, based on the achievement of specific objectives in listening/speaking. For this purpose, measures of individual achievement would be a secondary concern.

3. If possible, to provide teachers with models for their use in developing tests in listening/speaking based on specific objectives appropriate for their own students.

THEORETICAL FRAMEWORK

Oral communication involves both the reception and expression of ideas and information through spoken language. Effective oral communication not only conveys intended messages and meanings; it is also appropriate to the communication situation, the audience, and the speaker's purpose.

The assessment of functional communication competence and effectiveness is of current concern throughout North America. In a recently published book, Larson et. al. (1978) point out that functional communication refers to the skills, knowledge, and attitudes possessed by an individual, from which that individual's competence or effectiveness may be inferred. The lowest level of functional communication is competence. Competence in communication is behavior that meets the minimal functional demands of a situation and is socially acceptable, and most adults generally exhibit communication competence.

Larson et. al. further note that the levels of functional communication, above competence, imply degrees of communication effectiveness. Where competence involves minimal levels of communicative functioning, effectiveness involves communicating in such a way that certain desirable outcomes are enhanced or facilitated. However, a person may be an effective communicator in many different ways. As Larson et. al. point out:

> If another person understands the content of your communication or understands your unique personal

> characteristics, you may be effective in the sense that you have facilitated or enhanced accurate perception between you and the other. If you are able to change or modify another person's point of view on an issue, you may be considered effective in the sense that you have facilitated or influenced an outcome. . . . Many different outcomes are used both to conceptualize and to measure an individual's level of communication effectiveness. For the moment, the point is simply that communication effectiveness implies something more than, and different from, communication competence. (p.3)

Larson et. al. also suggest that the difference between competence and effectiveness is paralleled generally by different measurement strategies. They note that competence is usually assessed through objective-referenced or criterion-referenced tests, whose purpose is to determine whether or not a person demonstrates a specific ability or performs a particular task. The test score indicates whether or not the person taking the test responds correctly or adequately to the test items. Such tests are appropriate when examining the capability of students to perform each of a variety of communication operations. This type of assessment is different from norm-referenced testing, the purpose of which is to determine how well the student performs certain communication operations. In such testing, the score indicates how well the student can demonstrate a skill level compared to how well others being assessed can demonstrate this skill level. Norm-referenced testing is not concerned with the presence or absence of a particular communication knowledge or skill. Since the purpose of the MACOSA project was to develop test instruments in listening/speaking that could provide for program evaluation through the collection of baseline data about general student functioning, as well as to provide for diagnostic assessment of individual students rather than for comparison among students, it appeared reasonable to develop criterion-referenced tests which would provide evidence relative to whether or not students demonstrated specific abilities--not how well they demonstrated such abilities in comparison with each other.

While the importance of oral communication is accepted by educators considerably greater emphasis at all grade levels has been placed on the teaching of written reception and expression. As a result, little attention has been given to defining the objectives of listening/speaking within language arts programs at the elementary and secondary school levels. For example, Landry (1969) points out that "Although research has consistently shown that pupils spend more time in listening than in any other language arts activity, a serious lack of programs which develop listening skills is evident in most elementary schools." Devine (1978), in a study of listening research and theory over the past fifty years, notes that "Listening . . . is not generally accepted as part of the standard school curriculum. Surveys of actual classroom practices indicate little time devoted to direct instruction in listening skills, even though the evidence indicates strongly that such instruction works."

The importance of providing instruction in listening and speaking is echoed over and over in the literature; the lack of such instruction is echoed as well. In a report of a 1968 NDEA Speech Arts Institute (Vogelsang, 1972), the authors note that, at a conservative estimate, listening and speaking comprise between seventy-five and ninety percent of communication. They point out, however, that about seventy-five percent of educational training is directed toward the teaching of written communication--reading and writing. The dearth of test instruments in the areas of listening and speaking, as compared with the overwhelming number in reading and writing, attests to this lack of attention.

Examination of curriculum handbooks and guides developed by various Alberta school divisions reveals that most include some oral communication objectives. While these vary in number and specificity, they are usually concerned with oral reception and expression in formal rather than in functional communication situations. Since emphasis is generally on the use of language to inform, other uses of language are often dealt with minimally or totally neglected. The predominance of written language in school curricula is reflected in the fact that listening objectives appear to be mirror images of reading objectives and speaking objectives appear to be mirror images of writing objectives. Thus concern in listening is generally with those skills involved in comprehending written language read aloud rather than in understanding spoken language, and concern in speaking is generally with those skills involved in reporting and oral interpretation rather than with those skills involved in functional and interpersonal oral communication. Nevertheless, current curriculum development in Alberta places strong emphasis on the development of functional oral communication skills within a balanced language arts program. Thus the development of appropriate oral communication objectives and assessment devices is an urgent task in Alberta at this time.

Several of the recommendations contained in two reviews of language arts literature developed for MACOSA in 1977 emphasize the importance of developing objectives and assessment instruments that examine oral communication in functional situations for various audiences and purposes (Marchak, 1979; Plattor, 1977). These documents further reinforce the importance of assessment in classroom-type situations. Among these recommendations are the following:

1. Since oral reception and expression occur in situational contexts, it is most reasonable to measure these types of expression in such contexts. Therefore, controlled but natural testing situations should be used in assessment.
2. Reception and expression of spoken rather than written language should be a major concern in the construction of oral language assessment instruments.

3. Oral expression (speaking competence) should be tested in individual or small group situations.
4. Assessment of oral reception and expression should focus on the elements of spoken language that make it different from written language.
5. Since the communication of ideas and information involves both cognitive and affective aspects, both of these should be reflected in oral communication objectives and in assessment instruments designed to test oral communication.

In a recent paper (1978), Nancy Mead of the National Assessment of Educational Progress (NAEP) provides some additional recommendations for developing oral communication tests. (These recommendations are based on NAEP's experiences in attempting to develop listening tests.) They include:

1. Focus on skills which are unique and central to listening (and, by implication, to speaking). Mead notes that these skills involve responding to oral language, which is different from spoken language in that it tends to be nonlinear, incomplete, and redundant. However, she also notes that "It is less obvious how to deal with the nonverbal and interactive nature of oral communication in an assessment situation. Nonverbal signals tend to be subtle and individualistic and thus difficult to include in an assessment. Likewise, the give and take of normal speaking and listening are difficult to recreate in a test setting." (p. 10)
2. Use short, interesting, listening stimuli. As Mead notes, by focusing on topics and listening situations which are common to most students (for example, school activities, friendships, television), it is possible to provide stimuli that are both natural and interesting.

3. Consider extraneous factors which might contribute to item bias, e.g., the quality of the listening stimulus, the language level used, etc.

It is interesting to note the disclaimer following these recommendations:

> The recommendations discussed above are general and suggestive . . . . The recommendations emerge more from the subjective experiences of the NAEP/SCA pilot listening assessment project than from concrete empirical findings. These guidelines must be subjected to careful study and research . . . . (p. 14)

Nevertheless, the investigators in the MACOSA speaking/listening assessment development project used these as guidelines both in designing the tests and in drawing conclusions based on the pilot and field test data.

The concern with "natural spoken language" is basic to assessment of listening and speaking. It would appear that there are two main types of spoken language--spontaneous and prepared. Spontaneous spoken language is generated as the speaker talks and is organized as the speaker's thoughts take shape. Wilkinson et. al. (1974) mention two kinds of spontaneous spoken language: "old organized speech"--the fluent speech used in ordinary conversation in which the probability of certain words occurring together is high, and "new now organizing speech"-- the more hesitant speech used when speakers are thinking and choosing words as they go along. This speech is used to deal with cognitive problems requiring accurate solutions and therefore contains more redundancies, pauses, and false starts. On the other hand, prepared spoken language is organized prior to its use in a communication situation (as in giving a report from notes) and is generally in accord with a pre-determined purpose and audience. Both spontaneous and prepared spoken language may reflect casual, informal, and formal levels of language use.

However, both these forms of spoken language are different from written language in several ways. In written language, sentences are carefully structured and linked. Written language is tightly organized and contains more objective details and fewer ambiguous ones than does spoken language. Spoken language is often loosely organized and may be characterized by incomplete or ungrammatical utterances and repetitive false starts. Written language contains punctuation to indicate vocal factors such as pitch, stress, juncture, etc. Thus an oral reader must follow the writer's vocal clues, whether or not he/she thinks these are appropriate or feels that the listeners will understand them. Vocabulary and writing conventions must also be used to provide information given by the speaker through vocal and nonverbal factors.

It would appear, then, that emphasis in oral communication assessment should be placed on both spontaneous and prepared oral language. However, since students must often deal with written material read aloud, it is necessary to assess competence in oral interpretation and listening to written language as well.

The results of Loban's longitudinal study of children's language are also of value in that they provide recommendations for assessing oral communication. (The results were reported in two monographs, written at various phases of the study. One was published in 1963 and the other in 1976.)

Loban's findings were reported in four categories: (a) fluency--readiness and smoothness of speech; (b) effectiveness and control of language; (c) functions accomplished by the use of language; and (d) interrelationships among the various aspects of language. Among the findings are the following: (a) power over language was found to increase through the successive development of control over meaningful forms--ability to handle pronouns, grammatical subordination, etc.; (b) no basic ability differences appeared among various ethnic groups; (c) subjects more proficient in language showed more sensitivity to its conventions and more frequently used language to express tentativeness; and (d) there was a positive relationship among the skills of listening, speaking, reading, and writing, suggesting that competencies in these areas reinforce each other. Among Loban's conclusions are these: (a) to achieve language flexibility pupils must apply whatever is studied to situations in which they have something to say, a deep desire to say it, and someone to whom they genuinely want to say it; (b) situation plays a large part in language fluency; (c) linguistic growth stages parallel stages of physical growth; and (d) more research on language development and use is needed.

Based on Loban's conclusions and recommendations (1976), it appears that oral communication assessments must take into account motivation and situation, as well as the physical, intellectual, and linguistic ages and stages of students. This is consistent with the recommendations of other researchers in the field of oral communication.

In summary, research reviews suggest the importance of the following considerations in assessing oral reception and expression:

1. Objectives for listening/speaking should be specified in terms of knowledges, skills, and attitudes; audience, situation, and purpose; cognition and affect; casual, informal, and formal language use.
2. Objectives should be concerned with spontaneous spoken language, prepared spoken language, and written language read aloud.
3. Assessment devices should be concerned with knowledges, skills, and attitudes; stimuli should be short, varied, and concerned with different audiences, situations, and

purposes; casual, informal, and formal language use should be reflected in these stimuli.

4. Assessment devices should provide for the evaluation of expression and reception of spontaneous spoken language prepared spoken language, and written language read aloud.

5. Assessment devices should provide for the evaluation of a variety of oral language factors, including vocal, non-verbal, and verbal ones (vocabulary, syntax, etc.)

6. Assessment of oral communication competency should be conducted under controlled but natural testing conditions.

7. Attention should be given to examining available oral language test instruments to determine whether or not they meet the above criteria.

LIMITATIONS OF THE ASSESSMENT

The pilot test subjects involved in this study were limited to two intact rural and two intact urban classes at each of Grades 3, 6, 9, and 12. The field test subjects were limited to two intact rural and seven intact urban classes at the Grade 3 level; three intact rural and six intact urban classes at the Grade 6 level; four intact rural and seven intact urban classes at the Grade 9 level; and two intact rural and four intact urban classes at the Grade 12 level. The field test sample was chosen in order to provide approximately 200 students at each grade level.

ORGANIZATION OF THE STUDY (WORK PLAN)

The organization of the study can be more clearly stated in terms of A Work Plan comprising nine activities. These activities may be summarized as follows:

Activity 1 involved the preparation of lists of measurable objectives for listening and speaking based on the prioritized and validated objectives provided to the investigators by the Steering Committee. After reviewing the lists of objectives presented to the investigators by the Steering Committee (the product of an initial study commissioned by MACOSA and carried out by the Edmonton Public School District), it was agreed by the Steering Committee and the investigators that additional work would have to be done in order to produce appropriate lists of measurable objectives. Criteria which had been listed by the investigators in their Proposal were applied, and, where these criteria were not met, objectives were revised and/or added. These criteria included the following: (a) objectives must be significant in terms of teacher and student instructional needs in the areas of listening and speaking; (b) objectives must contain measurable attitudes, knowledges or performance (skills) outcomes; (c) objectives must include standards of success that are clearly defined. Thus the refinement process was a major concern of the Research Team during this activity.

Activity 2 concerned the development of a list of persons to serve as an Advisory Panel to the Research Team. It was initially proposed that the Panel members would be asked to propose items or item sources for the development of an item pool and that they would be asked to interpret the findings of the field testing and provide recommendations for the design of the final test instrument. However, it was later agreed to forego this activity for two main reasons: (a) time constraints in the development of this Project did not permit the structuring of the Panel in such a way that it would have been useful; and (b) expenditures in the field testing were such that monies needed to be reallocated from this fund to the cost of Research Assistants. However, on-going advice was solicited and received from colleagues of the investigators throughout the project, and this advice was very useful in developing the objectives and items.

Activity 3 involved the undertaking by the Research Team of a review of related literature in listening and speaking in order to validate its conception of the categorization of related behaviors and instructional domains in listening/speaking as well as to analyze existing measurement devices. The latter task was undertaken in order to locate and evaluate possible behavioral indicators which might be considered in measuring listening/speaking attitudes, knowledges, and skills; and to identify and evaluate measurement devices and items in terms of objectives, format, appropriateness to designated grade level, validity, reliability, factor structure, and administrative feasibility. The results of this task are discussed in Chapter 2 of this Report; however, it should be noted here that an extensive ERIC search and library review of the literature revealed very few appropriate test instruments for modification or use in this project.

Activity 4 involved the completion of item matrices for each test, test blueprints for each test, and domain classifications for each test. These were approved by the Steering Committee for use in pilot testing.

Activity 5 involved the construction of an item pool comprising potential items for the measurement of each of the prioritized objectives. In general, stimulus items and test items were constructed by the Research Team, as appropriate ones to test the prioritized objectives did not appear to be available in existing tests. The Research Team considered the possibility of adapting the Listening Comprehension Tests (Wilkinson, et. al., 1974) but later decided against this since the subtests of which this instrument is comprised contain a large number of items to test specific objectives and the decision of the Research Team and Steering Committee was to use approximately three items to test each objective.

Activity 6 consisted of the development and administration of pilot tests comprising selected items developed during Activity 5. Analyses of the data gathered during the pilot testing were undertaken. Pilot testing was carried out in May, 1978, by members of the Research Team in schools in Calgary, Okotoks, Airdrie, and Springbank. Discussions concerning reactions to the test instruments and administration were held with pilot school administrators, teachers, and students.

Activity 7 involved the development of a Sampling Plan for field testing in consultation with the Steering Committee, the evaluation, revision, and addition of items based on pilot test data analyses; the development of field test instruments and Administrative Manuals; orientation training sessions with participating field test teachers; and the administration of the field tests under the supervision of the Research Team.

The pilot tests in both listening and speaking were administered by certified classroom teachers hired and trained by the Research Team. The decision to involve such personnel rather than classroom teachers was made for two reasons: (a) to ensure that the oral communication evaluations would be as unbiased as possible; and (b) to prevent using an inordinate amount of the classroom teachers' time for test administration and scoring. (The administration procedures are described in greater detail in Chapters 3 and 4.)

The field tests in listening, the written subtests of the speaking tests, and the attitude tests were administered by classroom teachers. The oral subtests of the speaking tests were administered by certified substitute teachers hired and trained by the Research Team. Although it had originally been the plan of the investigators that the classroom teachers would do all of the actual administration, the pilot testing indicated that this would not be feasible. While the original Sampling Plan approved by the Steering Committee called for the testing to be done in Calgary and Didsbury, administrative arrangements with the Calgary Public School Board could not be finalized, due to the reluctance of teachers and administrators in the Calgary schools to participate in the project so late in the school year (mid-June). However, the Red Deer Public School District #104 was amenable to involving itself in this project, and so the urban population was taken from Red Deer.

During Activity 7, detailed preparation for the administration of the field test instruments was made in consultation with the selected schools and principals. Test administrators were trained by the Research Team. The tests were printed, packaged, and delivered to the sample schools. Audio-visual equipment was prepared as well. Tape recordings of the stimulus materials for the listening tests were made through the auspices of the Learning Technology Unit of the Faculty of Education, University of Calgary. Testing of Grades 3, 6, and 9 was completed in June, 1978. Because of the pressures of a province-wide English testing program, school-wide testing programs, and graduation activities, testing of Grade 12 was postponed until September, 1978. Tape recordings of the students' talks during the field test administration of the speaking tests were made by the test administrators. All tape recordings (listening/speaking), tests, and answer sheets were collected by the administrators and returned by them to the Research Team. (The administration procedures are described in further detail in Chapter 3.)

Activity 8 involved the production of a Data Analysis Plan, sets of scoring procedures, and a Reporting Plan. These were approved by the Steering Committee. This activity also comprised analyses of the data obtained in the field testing, as well as the evaluation, revision and addition of items for the final test instruments based on the field test data analyses.

Activity 9 consisted of the production of a Technical Report, a Condensed Report and the Final Test Instruments for listening/speaking for Grades 3, 6, 9 and 12. As a preliminary step in the production of the Technical Report and Final Test Instruments, an outline of each of these was presented to the Steering Committee. Upon approval, drafts of the Technical and Condensed Reports were written, and the Final Test Instruments were completed.

As outlined in the Proposal, the Technical Report includes conclusions and recommendations based on (a) the Research Team's evaluation of pupil strengths and weaknesses in terms of the objectives; (b) the Research Team's opinions as to which objectives require attention in terms of curriculum development and/or revision; (c) the Research Team's evaluation of procedures used in developing the listening and speaking tests as well as of the comprehensiveness of these procedures in relation to the objectives; (d) the Research Team's opinions about further testing based on the results of the pilot and field testing involved in this study; and (e) a statement of possible hypotheses that might account for strengths and weaknesses in student achievement in listening and speaking, as well as suggestions for possible studies that might be undertaken to test these hypotheses. This statement also includes recommendations for specific instructional techniques and innovations in the teaching of speaking and listening.

A draft of the Technical Report was submitted to the Steering Committee for reaction, and final revisions based on these reactions were made. Camera-ready copies of the Technical Report and a Condensed Report were submitted in completion of the study.

ORGANIZATION OF THE REPORT

This Report is divided into five chapters. Chapter 1 presents an introduction and overview of the study. Chapter 2 provides a brief review of pertinent literature in listening and speaking. Chapter 3 describes the design of the study. Chapter 4 reports the results of the pilot testing in listening and speaking, the revisions made in the pilot test instruments in the development of the field tests, the results of the field testing in listening and speaking, and the revisions made in the field test instruments in the development of the final test instruments. Chapter 5 presents conclusions and recommendations based on the study. Appendices follow Chapter 5.

Chapter 2

RELATED LITERATURE AND ITS IMPLICATIONS FOR TEST DEVELOPMENT

In Chapter 1 of this Report, an outline is presented of the major tasks undertaken by the Research Team in the preparation of tests designed to assess the level of achievement in listening and speaking of Alberta students at Grades 3, 6, 9, and 12. As described in this outline, a major task involved a review of literature in listening, speaking, and communication appropriate to this study, in order to validate the Research Team's conception of behaviors and instructional domains in listening/speaking, as well as to analyze existing instrument devices. This task is described as follows in the Proposal made by the Research Team to MACOSA (page 11):

> ACTIVITY 3: By March 15, 1978, the Research Team will complete a review of related literature and available measurement devices in listening/speaking.
>
> 3.1: The Research Team will undertake a review of related literature in listening/speaking in order to validate its conception of such matters as the organization of related behaviors and instructional domains.
>
> This review will include at least the following sources:
>
> (a) The Marchak (1979) and Plattor (1977) Reports
> (b) published research reviews
> (c) articles and other materials selected through ERIC
> (d) dissertation abstracts
> (e) psychological abstracts
> (f) published reviews of measurement devices.

Several research reports were commissioned by MACOSA prior to the inception of this study. The Marchak Report (1979) examined procedures and instruments used in assessing communication skills and provided a number of recommendations for the development of instruments designed to test listening and speaking. The Plattor Report (1977) examined selected research reports in the area of listening and also made recommendations for test development. Several of the recommendations from both of these Reports are summarized in Chapter 1 (see pages 52-53. In addition, an extensive section on the theoretical framework underlying this study is presented in Chapter 1 largely based on the review of the literature undertaken for this study. Thus the reader is referred to the Marchak and Plattor Reports (see Bibliography) and to the section titled "Theoretical Framework" in Chapter 1 of this Report for additional information gleaned from the review of the literature.

The literature review was begun with an extensive search of ERIC documents and those of Psychological Abstracts, through the University of Calgary's Computer Services. Many curriculum guides, professional textbooks, Teacher's Guides accompanying textbook series in language arts areas, and student texts were examined in an effort to determine listening/speaking objectives and curriculum materials currently in use in North America and, where appropriate, in England. Materials concerned with listening, speaking, and communication that were published by professional educational associations (e.g.: National Council of Teachers of English, International Society for General Semantics, Speech Communication Association, etc.) were also examined. Published review of measurement devices as well as other published materials concerned with the assessment of communication, speaking, listening, and attitude instruments were reviewed.

We examined a great many materials--more than can possibly be discussed within the limitation of time and space provided in this

Report. Therefore, we will describe in general terms the information obtained from the literature review, with special emphasis on assessment of instruments designed to test listening and speaking at the elementary and secondary grade levels. An extensive Bibliography is presented at the end of this Report, and the reader is invited to peruse these documents at his or her leisure.

FUNCTIONAL COMMUNICATION

A considerable amount of energy has been expended during the last decades in an attempt to develop and delimit the area of functional communication. As Larson et. al. (1978) point out, concerns with the educational process in North America have resulted in a number of efforts designed to deal with the problems found in student learning levels. Such efforts have generally focused on identification, development, and remediation of skills which will help students function more competently in society. Among these skills are those of communication, with special current emphasis on functional oral communication.

The writings of a number of philosophers, linguists, psychologists, and educators, both in North America and England, have established the importance of concern with functional communication in the teaching of language arts. While it is not the purpose of this review to provide references in this area--this is done extremely well in Larson et. al.'s new book, Assessing Functional Communication (1978)--it is important to note that most of the newer curriculum guides, professional texts, and student textbooks in language arts reflect the growing concern with those aspects of functional communication which are deemed most essential in the development of oral expression and reception. Such concern is reflected in discussions of teaching at all instructional levels and for most educational purposes. In the area of teacher education, for example, Hurt et. al. (1978) discuss the transactional nature of communication; the importance of audience, situation, and purpose in the communication act; and the multidimensional nature of human communication as it functions in the classroom. Klein (1977) further reinforces this notion in describing the "four factors of circumstance--setting, subject, audience, and purpose." (p. 16) Using setting as the criterion, he sets up a number of "factors of circumstance" in an effort to build a talk model for use in the classroom.

A number of recently published ERIC documents are concerned with the development of functional communication in various settings and at various instructional levels. Among these are Book and Galvin, Instruction in and About Small Group Discussion (1975); Wiemann and Wiemann, Nonverbal Communication in the Elementary Classroom (1975); and Wood (ed.), Development of Functional Communication Competencies, Grades 7-12 (1977). These phamplets are easy to read and apply, and deserve attention by anyone interested in the role of functional communication in teaching.

Several recent publications by National Textbook Company of Skokie, Illinois have added much information to the realm of functional communication. Among these are Allen and Brown (eds.), Developing Communication Competence in Children (1976) and Holtzman and Ecroyd, Communication Concepts and Models (1976).

Concern with the various types of language used in functional communication is basic to, and continuously underlies, the work of Wilkinson, Stratta and Dudley (1974). While their work is discussed earlier in Chapter 1 of this Report, it should be noted here that their contributions to the field of oracy and its applications to their listening test battery, the Listening Comprehension Tests (discussed later in this Chapter) have been basic to the development of the MACOSA listening and speaking tests. Wilkinson and his colleagues approach the study of spoken language in terms of this question: "Who communicates what to whom, how and why, and on what occasion?" They feel that all of the elements of this question must be considered when assessing language. Thus they take the point of view that communicative competence must be tested in terms of context (situation), audience, and purpose. They note that, in dealing with spoken language, teachers must ask what seem to be appropriate questions, and that this is the instructional model upon which teaching and language should focus.

The decision to consider knowledges, skills, and attitudes as well as various audiences, situations, and purposes in the MACOSA testing, then, has been based on the work of writers in the areas of thought and language, language development, inter- and intrapersonal communication, verbal and non-verbal communication, and functional communication. Among the references explored in these areas were Vygotsky (1962), Britton (1967), Barnes and Britton (1969), Inhelder and Piaget (1969), Sommer (1969), DeVito (1970), Furth (1970), Galloway (1970), Keltner (1970), Cazden (1972), Galvin and Book (1972), Hopper (1973), Rosenfeld (1973), Anderson et. al. (1974), Halliday (1974), Goldberg and Larson (1975), White (1975), and Lundsteen (1976).

Listening

The review of the literature in listening was generally concerned with three areas: (a) defining knowledges, skills, and abilities in listening and the research that impinges upon these areas; (b) providing appropriate instruction in listening; and (c) assessing listening competencies. Certain research findings in listening, and recommendations based on these, are presented in the Plattor Report, provided to MACOSA in 1977. The ERIC search and additional literature reviews done in conjunction with this study extended and reinforced the findings and recommendations in the Plattor Report. Several of these are summarized below. (Those concerned specifically with listening test construction are discussed in the next section.)

1. Listening assessment should be concerned with evaluating the ability to listen to speech in terms of differences in vocabulary, phonology, communication units, and uses of mazes. It must be based on specific objectives for listening instruction in these areas.
2. Listening assessment should be concerned with measuring acoustic, linguistic, and semantic features that listeners may use to "track" messages. It must be based on specific objectives in these areas.
3. Listening assessment should be concerned with evaluating environmental, discrimination, and comprehension skills. It must be based on specific objectives in these areas.
4. Since listening comprehension occurs in situational contexts, it is most reasonable to measure comprehension of listening in these contexts.
5. Listening comprehension should be tested under varying conditions of natural speech situations.
6. Listening assessment should measure the ability to make grammatical as well as meaning predictions in terms of spoken materials.
7. In order to measure communicative competence in oral language, tests should account for context, audience, and purpose. No matter how difficult this task, it must be undertaken.

In 1940, Arndt and Husband wrote an article for the English Journal which they simply called "Listening." In this article they note:

> The obvious place to commence a defensive program to start the development of more conscious and discriminative listening processes is in the school. Listening plays as important a part in the classroom as it does elsewhere. Very largely, the efficiency of classroom listening determines the amount and character of the learning done by pupils. (p. 375)

In 1969, in an article on Elementary English, Landry points to the neglect of listening occurring in the schools of that day. In 1978, in an article entitled, "Listening: What Do We Know After Fifty Years of Research and Theorizing?", Devine comments that there is now enough literature in listening to be able to make some intelligent decisions about teaching and learning in this area and to shape a reasonable rationale for curriculum development and teaching-learning strategies in listening. He does not, however, say that these things have been done. Indeed, after fifty years, the questions that Devine poses in the article are still being asked and there still do not appear to be definitive answers to such questions as "What is listening?" "Can it be taught?" and "Has a listening scope and sequence been spelled out?" In developing listening tests, then, the information that is currently available can be used, but the tentativeness of this information must be kept in mind.

Devine notes that, although many teachers and curriculum specialists recognize that listening is important, it is still not generally accepted as part of the standard school curriculum. A review of listening objectives in curriculum guides, textbooks, and professional references reinforces this statement. While more and more listening objectives are appearing in these references, they are generally very broad and many have more to do with attending behavior and general language arts skills and activities than with specific listening knowledges and skills, and are rarely developmental--that is, the same skills and activities are outlined but with a bit more sophistication in their description. As Devine puts it:

> Is there a scope and sequence underlying existing listening programs? The answer is yes, there are some, but they are almost always predicated upon the notion that the listening process is comparable enough to reading so that a sequence of reading skills can simply be "translated" into listening skills. (p. 301)

Certainly, the Research Team found the above statement appropriate in their review of objectives and instructional activities in listening. For this reason, it was necessary for the Team to spend a considerable amount of time translating listening objectives found in the literature into measurable and specific objectives pertaining directly to listening skills and not to associated reading ones. It should also be noted that most of the objectives found in the literature were concerned with the reading aloud of material developed in written language, rather than concerned with listening to material presented in a spontaneous or planned spontaneous mode, as described by Wilkinson et. al. (1974).

Devine also tackles the question: "Is listening simply 'thinking'?" and reports that there are no clear-cut answers from research. However, he says:

> Listening, at its simplest, probably involves minimal thinking activity; in its higher forms--that is, in the listening skills most often taught in the few school programs that try to improve them--listening certainly involves a good deal of high level mental activity. (p. 303)

Thus, the Research Team felt that it was important to consider the relationship between listening and thinking in developing the listening/speaking tests, and to attempt to develop tests which did test listening rather than thinking.

A number of research studies have looked at the question, "Can listening be taught?" (This is highly relevant to the question, "Should listening be tested?") In general, such studies indicate that it can be taught, while public opinion reinforces the notion that it should be taught. (See, for example, Caffrey (1949), Hollow (1955), Lundsteen (1971), Fisher (1973), Taylor (1973), and Ammon, (1974).) The findings of a research study by Fawcett (1966) in this area are revealing. Her study was conducted with Grade four, five, and six students, and was designed to determine the effectiveness of teaching listening skills to them. Among her findings were these: (a) students who receive listening instruction evidence significant improvement in listening ability, whereas those students who do not receive such instruction do not; (b) boys and girls do not differ significantly in listening ability; and (c) children's ability to use reference material is significantly related to listening ability. While this study was conducted more than 10 years ago, current findings reinforce hers.

A number of researchers have developed models for teaching listening. Horrworth (1966), Lundsteen (1971), Taylor (1973) and Tutolo (1977) are among these. Tutolo's model is particularly important, since it includes acuity or hearing; discrimination; and literal, interpretational, and critical comprehension. (This model was used in the development of the MACOSA listening tests.)

The area of environmental sounds was included in the grade three test based on a hierarchy of listening skills presented by Weaver and Rutherford (1974). The skills are included in the

"Environmental Skills" category (p. 1149). Weaver and Rutherford place most of these skills in the Kindergarten-Grade 3 area.

In 1965, Brown analyzed the speech and listening content of selected pupil textbooks in language arts for Grades 3-6. He found that listening was stressed rarely--that it was emphasized in .63 percent of the lessons and .57 percent of the pages in the texts that he examined. Further, listening was not found to be emphasized as a distinct area for direct instruction. As Brown notes: "Some books give no more attention to listening than the suggestion, 'Listen courteously.'" (1967, p. 341) While current researchers have not examined actual textbook content in listening to any great degree, observation of the content of recent language arts textbooks at the same grade levels as those examined in Brown's study reveals more concern with the area of listening. In addition, a number of materials developed in the form of kits are now available which emphasize listening. The growth of listening centres in classrooms has also been marked in recent years.

Researchers have also looked at the types of listening that people do, and have classified these in such ways as active/passive (Barbara, 1975); social/serious (Barker, 1971); and total or holistic (Egan, 1970). Facilitators for and barriers to effective listening have also been discussed. (See, for example, Campbell, 1958, and Barker, 1971.) Concern with the listening process itself also appears in the literature, with emphasis on the listening process as it functions in communication events, the effects of the process on animal and human interaction, and the function or aural selectivity in the listening process.

As mentioned earlier, one aspect of the review of the literature for this study was concerned specifically with the assessment of listening competencies. Again, the literature review extended and reinforced the findings and recommendations in this area made in the Plattor Report. These are summarized below (pp. 35-39):

1. Listening assessments must be based on specified objectives for listening instruction.
2. Appropriately developed criterion-referenced tests are more suitable for assessing listening than are norm-referenced tests.
3. Currently available commercial standardized listening tests are inappropriate since they are not based on any real theory of listening comprehension and since results obtained from their administration do not permit accurate assessment of levels of listening skill.
4. In developing tests which measure specific listening comprehension skills, care must be taken to exclude factors which in fact measure other aspects--such as memory or inferencing.
5. The use of such techniques as "chunked" items, Shannon technique, and cloze procedures should be investigated in the development of communication tests.
6. Concerning the Wilkinson test battery (1974), work should begin as soon as possible to Canadianize this instrument, since it measures certain vital parts of listening.

A number of general testing references and references concerned with defining and preparing educational objectives were examined as well as references concerned specifically with listening assessment. While space precludes discussion of all of these, several deserve mention. Among these were Bloom (ed.), A Taxonomy of Educational

Objectives: Handbook I, the Cognitive Domain (1956); Lindvall, Defining Educational Objectives (1964); Purves, et. al., Common Sense and Testing in English (1975); and Popham, Criterion-Referenced Measurement (1978). (The latter provided components of criterion-measured test specifications which proved very useful to us in the construction of both the speaking and listening tests.)

The materials on item development available from the California Objectives Exchange (1972) were also examined in depth.

A review was made of currently available standardized and non-standardized tests, many developed in master's and doctoral studies and some by commercial publishers. Several of the most-often discussed tests are described below. In general, none of these tests were found to provide for the measurement of those objectives specified by the Research Team and the Steering Committee. Critiques in the literatures of commercially developed tests generally expressed doubt as to the validity and stability of the tests.

Perhaps the most extensive review of currently available listening tests is found in Larson et. al., Assessing Functional Communication (1978). This book looked at 90 instruments, and examined seven in the specific area of listening reception. Of these, two were concerned with assessing the listening of children under eight years of age, and two with assessing the listening of college students, thus making them inappropriate for our purposes. One other was a test of phonology, although presented in a comprehension format, again making it inappropriate. This left the two major commercially available tests, the Brown-Carlsen, Listening Comprehension Test (1955) and the STEP Listening Tests (1957). Specific concerns with the validity and reliability of both tests have been stated in the literature, and neither is applicable to the objectives of the MACOSA study. In addition, both tests concentrate on assessing comprehension of written language read aloud, and do not generally deal with spontaneous language.

It is interesting to note that the Listening Comprehension Tests developed by Wilkinson, Stratta, and Dudley in England (1974) are not discussed in Larson's book. They are described in detail, however, in Wilkinson, Stratta, and Dudley's book, The Quality of Listening (1974). There are three batteries in these tests: Battery A for use with 10-11 year olds; Battery B for use with 13-14 year olds; and Battery C for use with 17-18 year olds. Each Battery consists of a number of tests which are designed to measure comprehension of different spoken materials. The following tests are included:

1. tests of content, which are designed to measure the ability to follow and to understand a piece of fairly informal exposition;
2. tests of contextual constraint, which are designed to measure the ability to infer missing parts of a conversation from what is actually heard;
3. tests of phonology, which are designed to measure the ability to understand differences in meaning brought about by different emphases in speech;
4. tests of register, which are designed to measure the ability to detect the appropriateness of the spoken language used in various situations;
5. tests of relationship, which are designed to measure the ability to detect the relationships existing between people from the language they employ.

The following statement, taken from page 95 of this book, sums up the work done by Wilkinson et. al.

> Validity in the more general sense, of measuring what one sets out to measure, is the hardest of all. In most experimental work various criteria are available and those commonly used include external and internal examination results, school-based test scores and teachers' assessments, the last usually in the form of rank-ordering. None of these are available in our work on listening. In this respect, the work has to be regarded as exploratory

> to be regarded as exploratory rather than definitive, although any subsequent workers in the field will in fact have possible validation instruments available as a result of it. It would be unfortunate if this unavoidable limitation were taken to detract in any way from the strength of the case presented. Listening tends to be taken for granted by most people, although it is by definition one of the vital basic skills. It needs to be taken seriously. It can be analysed; it can, if necessary, be measured; and it certainly needs to be more fully understood.

It would appear that the construction of this battery of listening tests is based on current and well-researched theory in the area of communicative competence. However, the subtests provide many items for assessing each of a few listening objectives. While these objectives are generally consistent with those specified for the MACOSA study, considerable time and energy would be required to isolate the small number of items needed for assessing each objective in criterion-referenced tests such as those developed in this study. In addition, Canadianization and the development of Canadian norms would be required as well. Both time and financial constraints. precluded the undertaking of these tasks by the Research Team.

Speaking

In addition to literature reviewed and reported above, a search was made of literature related specifically to speaking (speech). The object of this search was: (a) to investigate the present status of speech education; (b) to discover trends which would inform the selection of objectives and of items to test those objectives; (c) to investigate testing methods and procedures and analyze existing measurement devices; and (d) to discover which attitudes towards communication were considered most important by authorities in the field and what instruments were available for the measurement of these attitudes.

With regard to the present status of speech education, textbooks and articles reviewed evidence an escalating concern for the study of human communication. Educators are focusing on those issues which give impetus to the serious study of this aspect of human behavior. Speaking and listening are no longer considered "frills" but are seen to be skills necessary to survival.

Representative of this constantly reiterated concern is the following statement by Clark et. al. (1971) in "The Learning Encounter: The Classroom as a Communications Workshop."

> The central issue here, then, is that virtually anything is possible in our world--including total annihilation: but our central source of control today, our remaining instrument for essential meaning in a crucial situation is still human communication. Is it not, therefore, the teacher's responsibility to train students to participate knowingly and intellectually as crucial elements of a communication system? (p. 4)

Sanson (1965) states this concern in these words:

> Speech Education . . . (because the business of life depends on speech) might be considered the most important subject in the curriculum--indeed the central subject, affecting and influencing all the rest. But this is not so. The attention it receives is partial and haphazard. (p.1)

Johnson (1946) expands the concept of speech education when he says:

> Unless people are taught to speak and to listen in such a way that they strive for real understanding of one another, no matter how much voice and diction, or spelling and punctuation have been improved--communication has not been taught. (p. 481)

Two conclusions relative to the testing of speaking emerge from this concern with communication: (a) the testing of speaking behaviors in the schools is a task of vital importance; and (b) this testing must take into account not only the assessment of skills, but of skills which derive from knowledges and basic attitudes towards communication.

In addition, two constantly recurring emphases appear in the literature of the last decade. First, no single aspect of the field has received more attention in recent years than has interpersonal communication. This emphasis is reflected not only in articles, studies, and texts devoted solely to a consideration of interpersonal communication, but also in those dealing with public speaking, oral interpretation, speech education and, in some instances, even with speech remediation.

The second emphasis concerns the utilization of concepts derived from Communication Theory as a basis for the development of instructional materials in speaking. A classic statement of the behavioristic theories of communication is to be found in Berlo (1960). Miller (1967) provides a concise definition of the communication process: "Communication has as its central interest those behavioral situations in which a source transmits a message to (a) receiver(s) with conscious intent to affect the latter's behavior." (p. 4)

Most of the communication models employed in the development of instructional materials include: (a) a speaker; (b) a listener; (c) channels--verbal and nonverbal; (d) a "storehouse" of past experiences, knowledges and skills; (e) a situational context; (f) a feedback loop and (g) a noise source. These aspects are reflected in the Flow Chart prepared and used by the Research Team to provide direction in selecting of objectives, items and evaluation procedures (see page 3-10).

Concerning the selection of objectives, an article by Kibler and Barker (1970) called "Stating Objectives in Speech Communication," contained useful information for the development of objectives relating specifically to oral communication performance. The authors state:

> Behavioral objectives describe measurable speech communication behavior in the student who has successfully completed a learning experience.
>
> Behavioral objectives (i.e. planning objectives) provide meaningful and precise criteria for making judgments about the performance level of students who have successfully completed specific speech communication experiences. (p. 32)

Although the latter parts of these definitions could not be presumed to apply to the sample being tested, the framing of objectives was guided by the factors of: (a) describing measurable communication behaviors and (b) providing meaningful and precise criteria for making judgments about the performance level of students.

The authors also note that "As teachers specify behavioral objectives, they must also remain sensitive to substantial changes in emphasis occurring in the speech field." (p. 39) The changes in emphasis noted above guided the Research Team in the selection of objectives.

Curriculum guides and textbooks were also reviewed for the selection of objectives and items. Two forms of oral discourse impromptu or extemporaneous speaking and oral reading, emerged from this review as the most usable vehicles with which to test the areas of spontaneous speech, spontaneous prepared speech, and written language read aloud.

Cayer et. al. (1971) state that impromptu speaking can challenge the ability of the student as no other form of speaking does and that if the student is able to handle this task the teacher can assume he has a clear, working knowledge of what is required in speech-making and delivery. Impromptu speaking, then, appeared to be a necessary component. However, the consideration that students would not have had prior experience in speech-making mitigated against using impromptu speaking as a task and thus a compromise was made. Students would be given a limited time in which to prepare. By definition this speech task would then take the form of extemporaneous speaking. Mudd and Sillars (1969) define an impromptu speech as one for which the student has made no immediate preparation, and an extemporaneous speech as one for which the ideas, organization, and supporting materials are thoroughly prepared in advance. The speech task in this area at the Grade 9 and 12 levels incorporated elements of both impromptu and extemporaneous speaking in that the student was given only a short preparation period and was supplied with a limited number of supporting materials.

The research of Templin (1957), de Hersch (1966) and Loban (1963) as cited by Klatt (1972) influenced the selection of a picture sequence as the stimulus item for eliciting spontaneous speech at the Grade 3 and 6 levels. All speech textbooks devoted considerable space to speech organization. Callaghan (1971) states that there is a need to devote more attention to assessing students' ability to perceive relationships among the sub-ideas which support the central idea of speech and that: "Our most suitable examining tool . . . lies in the scrambled outline." (p. 112) This consideration led to the inclusion of such a scrambled outline in the speaking (written) tests at the Grade 9 and 12 level.

For the testing of oral reading a considerable body of literature was available for examination. Most of this literature deals with the subject at a level of sophistication which goes beyond the requirements of the MACOSA tests. See (Geiger, 1963); (Armstrong and Brandes 1963); (Lee, 1965); (Campbell, 1966); (Parrish, 1966); (Bacon, 1966); (Gray, 1968); (Fernandez, 1969); (Mattingly and Grimes, 1970).

Textbooks concerned with the teaching of speech in elementary and high school generally contain chapters on "Reading Aloud;" "Oral Reading;" or "Oral Interpretation." These proved useful for purposes of test construction. Representative of these texts are: (Ecroyd, 1960); (Mulgrave, 1961); (Streeter, 1964); (Huckleberry and Strother, 1966); (Fessenden, et. al., 1968); and (Baird and Knower, 1968).

Relative to the nature of oral reading, Mulgrave (1961) says that the task of the oral reader is to interpret the thoughts of an author to an audience in such a way as to convey both the meaning and mood of a selection. Without an understanding of meaning and an appreciation of the mood being recreated, the oral reader cannot effectively accomplish the stated purpose of oral reading. Thus, writers in the field strongly emphasize the importance of preparation for oral reading. Consideration of this led to the construction of a written test at all levels which directed the students to the analysis of the selection to be read.

Literature appropriate to the oral reading task at the various grade levels was sought and criteria for the selection of stimulus materials were established.

A search of the literature revealed a dearth of testing instruments. No standardized tests were available which would measure the test objectives agreed upon by the Steering Committee and the Research Team.

Greene and Petty (1959) state that measurement and evaluation in language is complicated by the fact that expression is both a process and a product. There is need, then, for a critical analysis of underlying skill areas. They list 24 general and specific oral language skills. They state also that in teaching children to think and talk under conditions imposed by the presence of an audience, emphasis must be placed on: (a) development of pleasing voice qualities; (b) clear enunciation of words; (c) gracious manners; (d) avoiding common language errors; (e) careful choice of words; (f) selection and arrangement of ideas; and (g) skill in expressing ideas in exact words and effective sentences. The authors make the point that effective audience-speaker relationships are the result of the interplay of the essential elements of communication and that many of these elements are so lacking in definition that thus far relatively little has been done to develop objective means of testing them.

A format for the testing of the above mentioned skills was sought. Reid (1971) states that at present no paper-and-pencil test exist that will measure the effectiveness of a speaker or of an oral reader, and that we need to rely upon the judgment of the observer. He suggests that it is necessary to provide the observer with a criticism form to assist him in arriving at a well-balanced judgment regarding the various elements of the art of speaking or reading and that numerical values should be assigned to each item on the form.

Many such rating forms are to be found in the literature. These were examined prior to the construction of rating sheets to be used in the evaluation of oral speech. Items were selected or constructed which would measure stated objectives. As mentioned above, based on these findings, a pencil-and-paper, multiple choice format was selected to evaluate the knowledges which were presumed to underlie the skills being tested. As far as it could be discerned from an examination of evaluation procedures found in the literature, this was a departure from currently used formats. However, it should be noted that, at advanced levels, speech knowledges are often tested by means of the written examination. (See Douglas, 1958; Gruner, 1968; Knower, 1966; and Kelley, 1971.)

ATTITUDE

An examination of curriculum bulletins, scholarly journals in the field of communication, and recent speech texts reveals a growing awareness of the relationship between speech and personality and of the influence the speaker's attitude towards communication has on his or her listening/speaking effectiveness. Representative of this awareness are such statements as: "Speech behavior is neither separable from personality nor trainable apart from personality as a whole" (Phillips, 1970, p. 309); and: "When we set about to teach speech or to learn to speak, it is not the mechanics of the act which are important--it is far more the functions of speech as they relate to the development of the whole personality with which we are concerned" (Weaver, 1963, p. 8).

It was therefore thought important that some measure of attitude towards communication be included in the MACOSA listening/speaking test.

The term "attitude" was defined by the Research Team (See p. 3-22), and the following aspects of communication attitudes were established: (a) motivation; (b) willingness to communicate; and (c) image of self and others. Each of these aspects appears related to the degree of apprehension a speaker experiences in a given speaking situation.

McCroskey (1977) provides an excellent review of recent theory and research in oral communication apprehension in an article appearing in Human Communication Research, the journal of the International Communication Association.

Virtually all textbooks in the area of speech contain a chapter on communication apprehension--variously described as speech anxiety, reticence, stage fright or unwillingness to communicate. The consensus of opinion seems to be that the greatest barrier to interpersonal communication is the fear of communicating thoughts, ideas, or feelings to others. Extensive research has identified between 10 and 20 percent of all college students and adults as severe communication apprehensives. Expert opinion suggests that a similar percentage may exist in school populations.

In the light of the above findings it was considered important that this underlying attitude (communication anxiety) be tested. Thus a measurement device for this assessment was sought.

Larson et. al. (1978) discuss three major ways in which this concept may be measured: (a) self-reports (introspection); (b) psychological changes; and (c) observer evaluations. They state that: "Self-reports are the most widely used, for a variety of reasons, mostly pragmatic. They are the easiest method to employ and probably the easiest to analyze. They have high reliability, and their validity is acceptable, though it is based primarily on construct validity" (p. 265). Wheeless (1975) gives a theoretical basis for use of the self-report:

> The point here is that if a person understands that he _is_ apprehensive and _why_ he is apprehensive (fearful) then his own report of his own fear ought to be the most valid. Observer ratings and physiological indices probably do not assess this aspect of communication apprehension as directly. If a person thinks he is afraid he probably is. The ways a person cognitively processes his own psychological cues probably determines his fear levels rather than the physiological manifestations themselves (p. 262).

Bormann (1969) reinforces the above with the comment: "The empirical researcher investigating attitudes must rely upon an introspective report of the subject as to the presence or absence of an attitude percept and its intensity" (p. 372).

A number of instruments designed to measure communication attitudes were examined. Observer ratings were ruled out since judges' ratings of internal anxiety states of speakers have not been found to be consistently reliable. Certain types of speech behavior (rate, pitch, nonfluencies, etc.) may be observable, and also may be a reflection of inner states. However, these observations must be made by trained assessors. Further, many behaviors associated with communication apprehension are impossible to observe. As Phillips (1970) notes: "A human being may be able to mask a personality disturbance by controlling overt speech behavior" (p. 32). In such a case, the internal state of the speaker would not be evident in his or her performance.

Many of the tests examined were designed for an adult population. In addition, most were too lengthy to meet the time constraints of the MACOSA total test situation.

A final choice of two instruments developed by McCroskey (1977) was made. The Personal Report of Communication Fear (PRCF), which McCroskey states can be administered to children from kindergarten age on, was chosen for administration at the Grade 3 and 6 levels. The Personal Report of Communication Apprehension (PRCA) was selected for administration at the Grade 9 and 12 levels, based on McCroskey's statement that this scale may be used successfully with high school students.

Chapter 3

DESIGN AND IMPLEMENTATION OF THE STUDY

OVERVIEW

This study was designed to develop tests for assessing the achievement of selected objectives in listening and speaking at the Grade 3, 6, 9, and 12 levels. The design for the study was initially outlined in a Proposal to the Minister's Advisory Committee on Student Achievement (MACOSA). It was reviewed and finalized during meetings with the principal investigators (Research Team) and the MACOSA Steering Committee. During the study, the Research Team compiled and prioritized objectives in consultation with the Steering Committee; reviewed pertinent literature in listening/speaking; constructed and administered pilot and field tests in listening/speaking based on the compiled objectives; constructed final tests in listening/speaking for use with students at Grades 3, 6, 9, and 12 in Alberta, and developed a Technical Report for MACOSA describing this assessment study.

This chapter describes the design of the study in terms of rationale and steps taken by the Research Team in developing the pilot, field, and final listening, speaking, and attitude tests. Chapter 4 details the construction and data analyses of the pilot and field tests, the item revisions required, and the construction of the final test instruments. Chapter 5 provides conclusions and recommendations.

RATIONALE

The products required by the contract between the Research

Team and MACOSA, according to the terms of reference for this study, were tests for use in Alberta schools in assessing the achievement of objectives in listening/speaking at the Grade 3, 6, 9, and 12 levels. The following guidelines were given to the Research Team:

1. The tests will be appropriate for each of Grades 3, 6, 9, and 12 for each of speaking and listening, and will be based on objectives selected by the Steering Committee for each of these grades and on the empirical results of pilot and field testing activities.
2. The tests will be approximately one hour long for each target grade for each of listening and speaking. (This was later changed, with the approval of the Steering Committee, to "each target student" rather than "each target grade".)
3. The tests will be rated as "good" using criteria established by the Center for the Study of Evaluation (CSE), including criteria such as reliability, replicability, range of coverage, and score gradation where these apply.

These guidelines were affected by a number of factors relevant to the study design.

Prior to the inception of this study, MACOSA had contracted with the Edmonton Public School Board (EPSB) for the development of a set of listening/speaking objectives for elementary and secondary schools in Alberta. This contract did not require that the objectives be developed in measurable terms. The objectives developed by EPSB reflected the broad scope of the language arts curriculum, with special emphasis on listening/speaking. Thus the Research Team felt that it would be essential as a first step in the study design to analyze the objectives as presented by the Steering Committee (these were selected from the EPSB objectives, and were prioritized and validated by the Steering Committee and Language Arts Policy Committee of the Department of Education); to categorize the objectives in terms of attitudes, knowledges, and skills; to indicate their relationships within these categories to spontaneous spoken language, prepared spoken language, and written language read orally; and to extend these categories and relationships in order to account for varied audiences, situational contexts, and listener/speaker purposes. This analysis and categorization was particularly important since the Steering Committee, during meetings with the Research Team, agreed that both affective and cognitive aspects of listening/speaking should be represented in the tests; that stimuli represent typical classroom activities; that common stimuli be used in the tests across grade levels where feasible; and that tests reflect various uses and levels of language.

During the course of this study, the Alberta Department of Education was working toward the finalizing of a new language arts curriculum for students in elementary and junior high schools in Alberta. This curriculum was planned to include much broader areas of oral communication than did the curriculum currently operational at these grade levels. While the Research Team had some knowledge about this new curriculum, it did not know the specific aspects of listening/speaking to be included.

In addition, emphasis on written language (reading/writing) in Alberta schools had, to this point, rather effectively precluded much instructional time devoted to oral communication at these grade levels. Further, as is common throughout North America where oral communication is considered a strong part of the curriculum by individual teachers, schools, or systems, the current Alberta curriculum at all grade levels was more concerned with transactional and literary elements (report presentations, oral interpretation, etc.) than with spontaneous talk (conversation, discussion, brainstorming, etc.). Thus, the Research Team found themselves in the unusual position of designing tests to assess future knowledges and expectations in listening/speaking as well as current ones, and in the even more unusual position of assessing knowledge and skill development in areas with which teachers might well be lacking knowledge and skill themselves.

Time and financial constraints had a strong effect on the development of the pilot, field, and final test instruments. Although the contract with the Research Team was not finalized until February, 1978, it was necessary that the entire project be completed by the end of September, 1978, in order to meet the requirements of MACOSA's mandate from the Alberta government. Thus, a period of only three months was available to develop and administer test instruments before the close of the 1977-78 school year. Needless to say, the computers worked overtime, as did the Research Team.

It had been originally planned by the Research Team that they would administer the pilot tests and that classroom teachers would administer the field tests. However, when the guidelines were changed to one hour of testing time per student, it was possible for the Research Team to do what they had wanted to do from the inception of the study--test speaking under individual rather than whole-class conditions. This would also allow for the testing of both affective and cognitive aspects of talk. (Findings in the field of human communications strongly suggest that emotional needs of the students should be considered as highly relevant to the speech act.) In addition, the Research Team felt strongly that each speaker was entitled to an audience, and that without an audience the communication stituation would be somewhat less than natural. Thus it was decided to administer the oral speaking skill portion of the test battery in small groups, using controlled pictorial stimuli at the Grade 3 and 6 levels, and a specific topic at the Crade 9 and 12 levels.

However, the administrative plan for individual oral testing required a considerable amount of time on the part of classroom teachers. Further, it required space allotment outside of the individual classrooms, and substitute teachers to cover classes while the classroom teacher was out testing the small groups. In addition, the testing procedure required that the talks made by students be taped, so that they could be scored under both live and taped conditions. This was necessary because it seemed highly improbable that teachers could rank the necessary facets of talk during a one- or three-minute live presentation. Thus the scoring time required by teachers was lengthy, and the Steering Committee was concerned (quite justifiably) with teacher reaction to this time expenditure, especially toward the end of the school term. To add to this, the Research Team was concerned with the possible effect of bias, since teachers knew their students well and might rank them on what they felt the students could do rather than on what they did do during the testing situation.

It was therefore decided to involve a group of certified substitute teachers as examiners. These teachers administered all of the Pilot Tests and the Oral sections of the Speaking Field Tests, while classroom teachers administered the Listening Field Tests, the written sections of the Speaking Field Tests and the Attitude Tests.

These substitute teachers were trained by the Research Team. However, the cost involved in this procedure was much greater than had been anticipated by the Research Team, thus requiring a significant amount of money to be transferred from the allocation for an Advisory Panel to the test administration. However, consultants were available from knowledgeable colleagues of the Research Team.

Time and financial constraints also made it necessary to limit the sample in terms of geographic area. Thus pilot testing was done in Calgary and surrounding rural areas, while field testing was done in Red Deer and Didsbury.

In order to place greater emphasis on language production than on paper-and-pencil testing (an emphasis expected by the Steering Committee and approved by the Research Team), it was agreed that knowledges in listening/speaking and attitudes would be tested primarily by pencil-and-paper tests; that listening would be tested through taped stimulus materials and paper-and-pencil response devices; and that speaking skills would be tested with oral language production. There was also strong concern with the development of appropriate Administration Manuals, so that the testing and scoring procedures would be controlled.

The time, financial, and school situation constraints made it necessary to limit the range of communication events that could be tested. One area considered of vital importance, based on the review of the literature as well as the expertise of members of the Research Team, was that of small-group or interpersonal communication. It was agreed to eliminate this area from testing not only because of the time factor but also because measures of such communication tend to be sophisticated and their interpretation therefore requires skills which the teachers could not be presumed to possess. Further, communicative interchange in small groups, both qualitative and quantitative, is largely dependent on classroom climate, and tests of such interchange might not measure a student's speaking ability or self-motivation, but rather a teacher's ability to establish a classroom climate conducive to free communicative interchange.

In the area of listening, it was clear that the most effective way to develop tape recordings for testing was to use professional actors and tape under professional studio conditions. However, both time and financial constraints, as well as the knowledge that the tapes would have to be made several times in order to incorporate revisions discovered to be necessary during the pilot and field testing, precluded the use of professional studio time during the making of the pilot and field test tapes. Thus sound reproduction during the pilot and field testing was not always the best quality. (The audio equipment in some of the schools didn't help any either.) However test administrators did not consider this a problem. Furthermore, these conditions were remedied in the production of the final test instruments.

It should be noted that stimulus materials were often taped in actual classrooms to provide authenticity. Readings of prose and poetry were done by practicing teachers, and student talks were provided by actual students. No professional actors were used, as it was felt that all stimulus materials should represent as much as possible the classroom conditions and situations.

The guidelines presented to the Research Team called for consideration of the use of all or part of available and appropriate listening and speaking tests. In the area of speaking, a review of available test instruments indicated that none could comprehensively test the objectives developed by the Research Team and the Steering Committee. Thus all items, stimulus materials, and scoring devices were developed by the Research Team. (In some cases stimulus materials were published excerpts from stories, poems, or plays.)

Only one available listening test instrument appeared to have possibilities--the Listening Comprehension Tests developed by Wilkinson, Stratta, and Dudley (1974) under the auspices of the

Schools Council in England. However, the use of this instrument was rejected for several reasons: (a) time constraints made it impossible to Canadianize the test in time for the pilot and field testing; (b) the norm-referenced nature of this test made it necessary to provide a larger number of items to test the objectives on which it was based than was needed for the MACOSA tests, and time and financial constraints precluded the research necessary to determine which items could be used individually rather than using the battery as a whole; (c) certain of the objectives tested were not appropriate to the MACOSA tests, while others that were appropriate to the MACOSA tests were not tested by the Wilkirson battery. However, a recommendation was made by the Research Team that Canadianization of this test as well as norming of it with Alberta students should take place, since a Canadianized version of this battery appeared to be an appropriate complement to the MACOSA tests. (See Chapter 5.)

Since attitudes toward communication were considered to be important, tests that might be related to the attitude objectives were examined. For this purpose, the Research Team finally selected the Personal Report of Communication Fear (PRCF) (McCroskey, 1977) for administration to Grades 3 and 6. The Personal Report of Communication Apprehension (PRCA), by the same author, was used to test Grades 9 and 12.

DESIGN STEPS

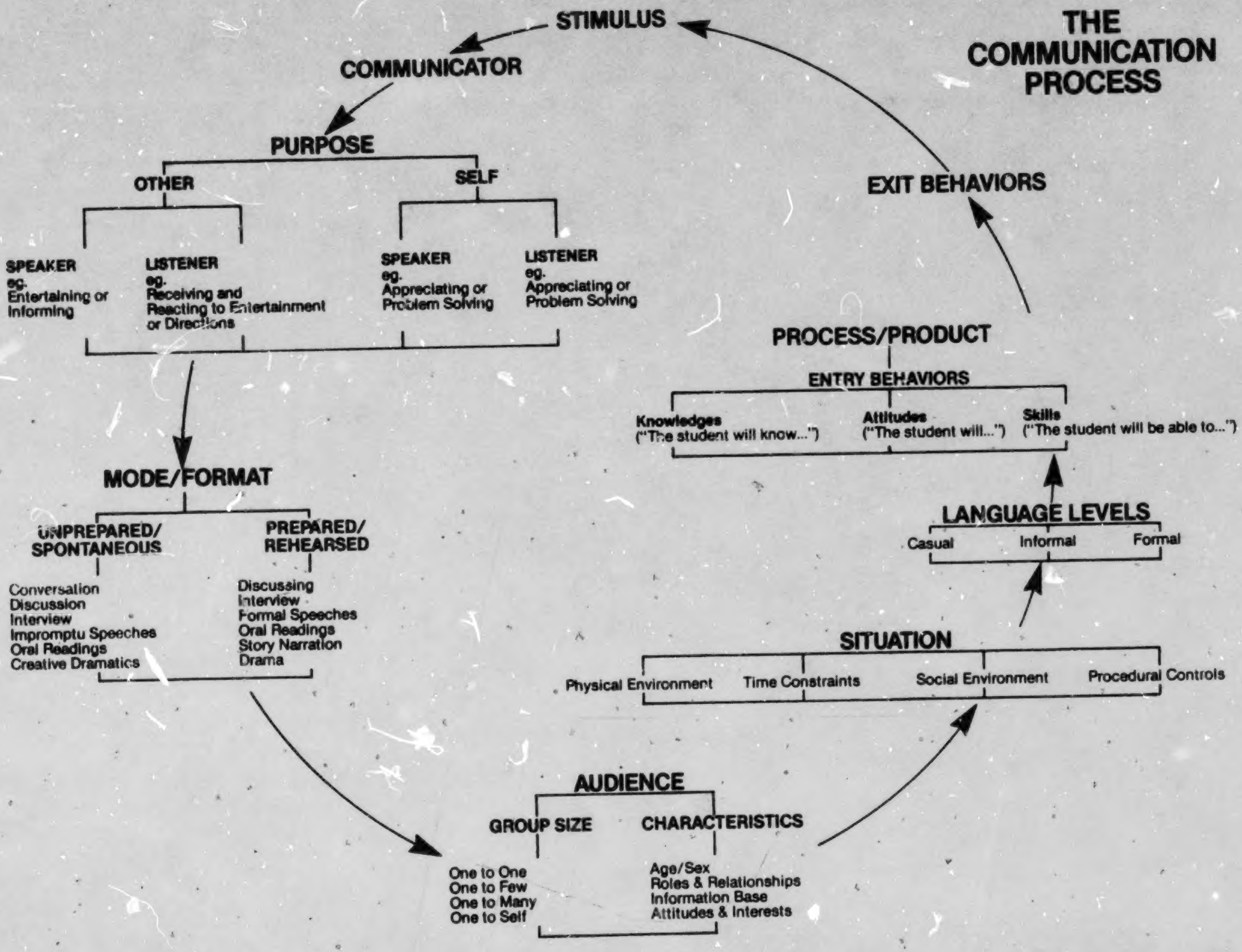
Step 1 The first activity in the Proposal made by the Research Team to MACOSA involved the preparation of relevant, high-priority, measurable objectives for listening/speaking. Therefore Step 1 involved the analysis of objectives presented to the Team by the Steering Committee as well as those which experts in the fields of communication and language arts consistently considered important. This analysis was done for several reasons: (a) to develop a list from which categorizations might be made in terms of attitudes, knowledges, and skills in listening/speaking; (b) to indicate the relationships of objectives within these categories to spontaneous, prepared, and written language read orally; and (c) to extend these categories and relationships in order to account for varied audiences, situational contexts, and listener/speaker purposes. This step also involved the matching of the lists developed by the Research Team with those presented by the Steering Committee.

Step 2 consisted of reviewing the listening/speaking/communication literature, and took place at the same time as Step 1. Thus findings could be immediately applied to the listing and categorization of measurable objectives. (See Chapter 2.)

Step 3 involved the structuring of the general domain of listening/speaking including knowledges, skills, and attitudes, so that clear links could be established among such factors as the communicator (listener/speaker), the purpose of the communication, the mode of communication, the situation, the stimulus, etc. The elements of this domain are diagrammed in Figure 3-1 on the next page.

From Figure 3-1, it can be seen that the communication process is a closed loop in which the communicator and the stimulus are the anchors. In this structure, the instructional objectives are seen as the products of the purpose. For example, we can say that a stimulus that creates arousal can be acted upon only to the extent that knowledge of audience, situation, modes of presentation, and purposes are taken into account in producing some form of output which is supported by specific attitudes.

The advantage of such a flow chart is that any specific objective can be related to any stimulus, audience, mode, purpose, and communicator (listener/speaker), and at any level of stimulus


STIMULUS
COMMUNICATOR
PURPOSE
OTHER
SELF
SPEAKER
eg.
Entertaining or Informing
LISTENER
eg.
Receiving and Reacting to Entertainment or Directions
SPEAKER
eg.
Appreciating or Problem Solving
LISTENER
eg.
Appreciating or Problem Solving
MODE/FORMAT
UNPREPARED/ SPONTANEOUS
Conversation
Discussion
Interview
Impromptu Speeches
Oral Readings
Creative Dramatics
PREPARED/ REHEARSED
Discussing
Interview
Formal Speeches
Oral Readings
Story Narration
Drama
AUDIENCE
GROUP SIZE
One to One
One to Few
One to Many
One to Self
CHARACTERISTICS
Age/Sex
Roles & Relationships
Information Base
Attitudes & Interests
SITUATION
Physical Environment
Time Constraints
Social Environment
Procedural Controls
LANGUAGE LEVELS
Casual
Informal
Formal
PROCESS/PRODUCT
ENTRY BEHAVIORS
Knowledges
("The student will know...")
Attitudes
("The student will...")
Skills
("The student will be able to...")
EXIT BEHAVIORS

THE COMMUNICATION PROCESS

Figure 1

and communicator sophistication.

Step 4 involved the devising of a set of indicators to be used in determining when students had received an acceptable level of attitude, knowledge, and skill sophistication. Some example indicators are:

(a) given a pictorial stimulus involving a sequence of events, the student will tell a story orally about the events depicted in the pictures;

(b) given a selection within the student's easy reading ability, the student will be able to read it aloud conveying the meaning and mood of a selection;

(c) given an oral report, the student will be able to recall sequences of details in the report; and

(d) given a selection from a conversation, the student will be able to indicate the effect of the intonation patterns of the speakers on the listeners.

Step 5 involved the selection and categorization of the final listening/speaking objectives by attitudes, knowledges, and skills. Further, it involved the placement of these categorized objectives into domains. Each objective was categorized, a priori, into a domain and a cluster. Approval was sought and received from the Steering Committee to use these categories for the initial design and evaluation of the test instrument. This list of objectives and domains is discussed later in this chapter under the heading "Objectives."

Several criteria were used by the Research Team for the selection of objectives. These included the following: (a) the objectives must be significant. It was the view of the Research Team that even well-accepted published tests often address objectives that are trivial or that teachers cannot do anything about; (b) the objectives must contain an observable attitude, knowledge, or performance (skill) outcome. The Research Team felt strongly about this criterion, because it had been their experience that objectives are often stated in terms that are not measurable and for which no indicators can therefore be determined; and (c) a standard of success must be defined in order for an objective to be significant. Thus, the objective that a pupil can understand a factual communication may indeed be significant, but indicators must be established that will clearly reveal the student is understanding. In addition, levels of competence must be established which take the teacher or examiner beyond being able to say that the student understands "quite a bit" or "50 per cent of the time."

Step 6 involved a re-evaluation of existing objective-based norm-referenced and criterion-referenced tests in communication, listening, and speaking. It further involved a consideration of their applicability to the objectives of this study. As noted earlier, this review revealed that none of the available tests could be used to assess comprehensively the finalized listening/speaking objectives. Therefore all of the items used in the Listening/Speaking Pilot and Field Tests were constructed by the Research Team. However, two attitude tests, the PRFC for Grades 3 and 6 and the PRCA for Grades 9 and 12, were found to be appropriate to the attitude objectives. Thus these instruments were used in their entirety, and administered as part of the speaking test battery.

Step 7 involved the construction of the Pilot Test item pool. The major purpose of the pilot testing was to get information on format, testing procedures, test length, and difficulty and discrimination indices for the items. In speaking, in particular, it was important to determine whether teachers were able to make the necessary discriminations on the rating forms, as well as to determine the reliability of these ratings. This phase of the study is discussed in greater detail later in this Chapter under the heading "Description of the Items."

Step 8 involved the development of Pilot Test Blueprints for the construction of the listening/speaking tests at each grade level. These contained a description of the domains and objectives used in each test, the number of items per objective, the number and per cent of objectives in the domain, and the number and per cent of items in the domain.

Step 9 involved the construction of the Pilot Test instruments. During this phase, the Research Team considered test formats and response modes, and decided on a group-response format for testing listening skills, listening/speaking knowledges (tested as part of the speaking test battery), and attitudes (also tested as part of the speaking test battery). They chose an individual-response format for testing speaking skills. Scoring devices, including checklists for ranking speaking competence, were developed. Student test booklets and administration manuals were written and printed. Tape recordings were prepared for presentation of listening stimulus materials and items. Visuals were developed as stimulus items for certain of the speaking tasks.

Step 10 involved the administration of the Pilot Test instruments. This was carried out in May, 1978. The Research Team selected the pilot sample, which included two intact rural and two intact urban classes at each of Grades 3, 6, 9, and 12. (This sample is described in greater detail in Chapter 4.) The administration was done by certified teachers (other than the sample students' classroom teachers) who were trained by the Research Team. Meetings were held with administrators and teachers in the sample schools to obtain cooperation and explain procedures. It should be noted that these schools included both closed and open area classrooms.

Step 11 was concerned with the analyses of the Pilot Test data. This involved the determination of item statistics, including difficulty and discrimination indices, choice distributions, reliability coefficients, and, for the speaking tests, rater reliabilities. This information was used in selecting items for the field tests, as well as for revising items where necessary. In general, items on written tests were deleted or revised if they were too easy (difficulty above .80); if they did not discriminate (Pt. Bis. below 20 on Stimulus Sub-Scale); or if options were chosen by fewer than five (5) per cent of the respondents. Items on rating scales were deleted or revised if the distribution of responses was obviously skewed or if raters were unable to complete them. Comments about the tests made by teachers, raters, and students were also taken into account. It should be noted that some items were retained even if they did not meet the Difficulty Index criterion of <.80. This was done because it was felt that a high priority objective was being measured in line with the basic concept of a criterion-referenced approach to testing. (See Chapter 4 for further discussion.)

Steps 1-11, then, were concerned with the identification and organization of the Pilot Test objectives, the construction of items, the development and administration of Pilot Test instruments, and the analyses of the data obtained from the pilot testing.

Step 12 involved determining appropriate revisions of objectives and items based on the pilot test data analyses and the constructing of additional items to replace those that were inappropriate or faulty. Descriptions of these revisions for both the speaking and listening tests at each grade level are presented later in this chapter.

Step 13 involved the development of Field Test Blueprints for the listening and speaking tests at each grade level. Like the Pilot Test Blueprints, the Field Test Blueprints contain a description of the domains and objectives used in each test; the number of items per objective; and the number and per cent of the objectives in the domain; and the number and per cent of items in the domain.

Step 14 comprised the construction of the Field Test instruments based on the results of the pilot testing. During this phase, stimulus materials were revised or replaced where necessary; items were revised, replaced, or added to where necessary; directions in student and administration manuals were clarified; scoring devices were revised in terms of pilot test data analyses and consultations with pilot test teachers and administrators; tape recordings were re-made; and student test booklets and administrative manuals were printed.

Step 15 involved the development, in consultation with the Steering Committee, of a Sampling Plan for the field testing. The first task in designing this Plan was a determination of the significant stratifying variables that had to be taken into account. Since it was the understanding of the Research Team that the major concern in this project was with the general competence levels of students in listening and speaking, it was recommended that the Field Test sample be composed of intact classes at each grade level, thus using schools rather than students as the sampling unit. The Research Team also felt that adequate field testing for data analysis purposes could be achieved through the use of approximately 200 subjects at each grade level. In addition, they felt that it would be useful to administer both the listening and speaking tests at the same time to the same population, both for greater test administration ease and for a better analysis base. All of these recommendations were accepted by the Steering Committee.

In light of time and financial constraints, it was agreed that the sample Field Test schools selected would be within a reasonable distance of Calgary as well as within Calgary itself. However, the Calgary Public School Board was unable to provide the requisite number of students at this time in the school year, so the urban sample was taken from Red Deer. This sample comprised seven intact classes in Red Deer at the Grade 3 level and two intact classes in Didsbury; six intact classes in Red Deer and three in Didsbury at the Grade 6 level; seven intact classes in Red Deer and four in Didsbury at the Grade 9 level; and four intact classes in Red Deer and two intact classes in Didsbury at the Grade 12 level. (The field test sample is further described in Chapter 4.)

Step 16 comprised the administration of the Field Test instruments. The Research Team delivered to each school copies of the student test booklets, administration manuals, checklists, etc. In each school, the principal was asked to act as coordinator for the testing program. (In some schools, the principal delegated this responsibility to another administrator or teacher.)

The testing of Grades 3, 6, and 9 was completed in June, 1978. The testing of Grade 12 was completed in September, 1978. (This later date resulted from a reluctance of school administrators to involve Grade 12 students in a testing program of this magnitude so close to graduation.) The administration of the Listening Tests, the Written subtests of the Speaking Tests, and the Attitude Tests was done by classroom teachers. The administration of the Oral subtests of the Speaking Tests was carried out by certified substitute teachers trained by the Research Team.

Step 17 involved the development of a Data Analysis Plan, which was presented to and approved by the Steering Committee. The Data Analysis Plan called for two approaches. While the basic approach to this study involved analyses appropriate to criterion-referenced testing--that is, determining acceptable levels of performance based on selected objectives--, it was decided that the usual procedures used for norm-referenced testing would also be applied. Thus the analyses involved deriving information based on item statistics and standard procedures for determining inter-item reliability. In addition, the Research Team performed factor analyses of all tests. The results of these analyses are presented in Chapter 4.

Step 18 involved the preparation of the Final Test instruments at the Grades 3, 6, 9, and 12 levels.

Step 19 involved the preparation of a Reporting Plan, the development of drafts of the Technical and Condensed Reports for submission to the Steering Committee for their consideration and approval.

Step 20 involved the development of the Technical and Condensed Reports. Conclusions and recommendations are included in Chapter 5 of this Report.

In summary, this study was concerned with assessing the achievement of selected objectives in listening and speaking at the Grade 3, 6, 9, and 12 levels, and was completed in a series of twenty steps.

Steps 1-11 were concerned with identifying and organizing the listening/speaking objectives, reviewing the literature, constructing Pilot Test items, constructing Pilot Test instruments, administering these instruments, and analyzing the data obtained from the test administration.

Steps 12-20 were concerned with revising stimulus materials and items where indicated by data analyses, constructing additional items where necessary, constructing Field Test instruments, administering Field Test instruments, analyzing the data obtained from the test administration, revising stimulus and pilot items where indicated by data analyses, constructing additional items where necessary, constructing final test instruments, and preparing the *Technical* and *Condensed Reports*.

OBJECTIVES

The first phase of this study was outlined in the *Proposal* as follows:

> *ACTIVITY 1: By February 28, 1978, the Research Team will prepare lists of measurable objectives for listening/speaking which we feel reflect MACOSA's requirements. These lists will be based on the prioritized and validated objectives provided to us by the Steering Committee.*
>
> 1.1: Between February 15-28, the Research Team will undertake an analysis of the proposed objectives for listening/speaking. The purpose of this analysis will be:
>
> (a) To categorize the objectives in terms of listening/speaking attitudes, knowledges, and skills;
>
> (b) to indicate their relationship within these categories to spontaneous spoken language, prepared spoken language and written language (read aloud), as these apply to communication tasks in both transactional and literary modes;
>
> (c) to extend these categories and relationships in order to account for varied audiences,

situational contexts and listener/speaker
purposes.

Thus, an initial set of prioritized objectives was presented to the Steering Committee and approved by them in March, 1978. (See Appendices A-1-4.) These objectives were classified by attitudes, knowledges, and skills. Justification for this procedure can be found in the literature, where oral communication is generally seen to contain both cognitive and affective elements, and where cognitive elements are usually seen to contain both knowledges and skills. Further justification can be found in the fact that courses in speech, at whatever level, contain both affective and cognitive elements, with specific concern directed to the development of both knowledges and applicational skills.

It was also hypothesized by the Research Team that a student might know the principles of message preparation and reception or of oral reading and yet not be able to apply them in the speaker-audience situation. If this were the case, it might be due to such aspects of lack of knowledge of factors involved in oral communication, lack of experience in oral performance, or fear in the speaking situation. Thus it was essential to isolate these various factors through evaluation of knowledges, skills, and attitudes in oral communication.

In a document entitled Specific and Prioritized Objectives for Listening and Speaking, prepared for the Steering Committee by the Research Team and presented to them on March 7, 1978, the sets of categories used to classify the knowledges, skills, and attitudes dimensions were enumerated and described. The following discussion is taken directly from this document.

Listening/Speaking Knowledges

Knowledge is defined here as the apprehension and comprehension of the elements of communication. It includes "knowing that" and "knowing how."

The Research Team felt that the same knowledges are necessary for effective communication on the part of both listeners and speakers. The speaker must use these knowledges to compose single and short utterances as well as longer pieces of talk. The listener must use these knowledges to comprehend appropriately these single or short utterances and longer pieces of talk--to "get the message" as the speaker intended it to be interpreted. Therefore the categories of knowledge objectives listed below apply to both listening and speaking.

Categories of Listening/Speaking Knowledge Objectives

1. Environmental Sound Factors.

2. Voice Production Factors.
   - A. Volume (Projection, stress, loudness/softness).
   - B. Pitch (High/low levels).
   - C. Juncture (Pausing).
   - D. Quality (Pleasant/unpleasant).
   - E. Rate and Fluency (Pacing, pausing and phrasing).

3. Nonverbal Factors.
   - A. Posture.
   - B. Body Movement and Gesture.
   - C. Audience Contact.
   - D. Facial Expression.

4. Linguistic Factors.
   - 4.1. General Linguistic Factors.
   - 4.2. Specific Linguistic Factors.
     - A. Sounds (Articulation and Pronunciation).
     - B. Words (Meaning and Structure).
     - C. Sentences (Meaning and Structure).

5. Organizational Factors.
   - A. Ideational Fluency,
   - B. Plan Analysis.
   - C. Organization of details.

6. Evaluation Factors.
   - A. Environmental Sound Factors.
   - B. Voice Production Factors.
   - C. Nonverbal Factors.
   - D. Linguistic Factors.
   - E. Organizational Factors.

Listening/Speaking Skills

A skill is defined here as the ability to perform an activity as measured by the end result of a physical, cognitive, or motor activity. Since the Research Team felt that there are differences in some aspects of listening and speaking skills, these were classified separately. The categories of skill objectives in listening and speaking are listed below.

Categories of Listening Skill Objectives

1. Environmental Sound Factors.
    - A. Identification.
    - B. Association.
2. Voice Production Factors.
    - A. Volume (Projection, stress, loudness/softness).
    - B. Pitch (High/low levels).
    - C. Juncture (Pausing).
    - D. Quality (Pleasant/unpleasant).
    - E. Rate and Fluency (Pacing, pausing and phrasing).
3. Nonverbal Factors.
    - A. Posture.
    - B. Body Movement and Gesture.
    - C. Audience Contact.
    - D. Facial Expression.
4. Linguistic Factors.
    - A. Sounds (Articulation and Pronunciation).
    - B. Words (Meaning and Structure).
    - C. Sentences (Meaning and Structure).
5. Organizational Factors.
    - A. Ideational Fluency.
    - B. Plan Analysis.
    - C. Comprehension of Organized Details.
        1. Literal Level.
        2. Interpretational Level.
        3. Critical Level.
6. Evaluation Factors.
    - A. Environmental Sound Factors.
    - B. Voice Production Factors.
    - C. Nonverbal Factors.
    - D. Linguistic Factors.
    - E. Organizational Factors.

Categories of Speaking Skill Objectives

1. Environmental Sound Factors.

2. Voice Production Factors.
    - A. Volume (Projection, stress, loudness/softness).
    - B. Pitch (High/low levels).
    - C. Juncture (Pausing).
    - D. Quality (Pleasant/unpleasant).
    - E. Rate and Fluency (Pacing, pausing and phrasing).

3. Nonverbal Factors.
    - A. Posture.
    - B. Body Movement and Gesture.
    - C. Facial Expression.

4. Linguistic Factors.
    - A. Sounds (Articulation and Pronunciation).
    - B. Words (Meaning and Structure).
    - C. Sentences (Meaning and Structure).

5. Organizational Factors.
    - A. Identional Fluency.
    - B. Plan Analysis.
    - C. Organization of details.

6. Evaluational Factors.
    - A. Environmental Sound Factors.
    - B. Voice Production Factors.
    - C. Nonverbal Factors.
    - D. Linguistic Factors.
    - E. Organizational Factors.

Attitudes

Attitude is defined here as an enduring learned disposition to behave in a consistent way toward a given class of objects. An attitude is measured by the consistency of response to that class of objects.

An attitude may be inferred from the motor expressions that accompany and indicate emotions. Emotion is defined here as a mental state which is inferred from motor expressions that accompany the state.

As with the knowledge objectives, the Research Team felt that the same attitudes are necessary for effective oral expression and reception. Therefore the categories of attitude objectives listed below apply to both listening and speaking.

Categories of Attitudes Objectives

1. Motivation.
2. Image of Self and Others.
3. Willingness to Communicate.

At this point, then, the first purpose of the objectives analysis had been met--to categorize the objectives in terms of listening/speaking attitudes, knowledges, and skills. It should be noted that other oral communication researchers and test developers have used other classifications to categorize listening and speaking objectives--such classifications are by nature arbitrary. However, it was felt that the classifications used in this study were appropriate in terms of the prioritized and validated objectives selected by the Research Team and the Steering Committee. Further, the classifications had the advantage of simplicity and generality and were therefore acceptable as listening and speaking domains for use in developing item matrices and test blueprints.

DOMAIN DEVELOPMENT FOR PILOT TESTING

Following the classification of the objectives, it was necessary to develop a set of listening and speaking domains, a procedure considered essential in facilitating the data analyses. As noted previously, the classifications into which the objectives had been placed were considered as the basis of the domain generation. Certain changes seemed appropriate, however, and these are described below. After consultation with the Steering Committee, the domains were finalized, and these served as the basis for the development of the Pilot Tests. With the domain finalization, the objectives were revised for a-priori placement within them. The total sets of listening and speaking Initial, Pilot, Field, and Final Test objectives, listed by domain, are found in Appendices B-1 and B-2.

At this time a very important decision was made concerning the objectives. Since it looked as if the Listening Tests would be quite lengthy because of the number of prioritized and validated objectives which required assessment, it was decided to confine the listening testing to skills, and to test oral communication knowledges wherever possible through the written elements of the Speaking Tests. Even with this consideration, several of the listening/speaking pilot subtests were found to be too long and were therefore reduced in length for field testing. The following listening and speaking domains were approved by the Steering Committee for pilot testing.

Listening

1. Environmental Sound Factors
2. Voice Production Factors
3. Linguistic Factors: Words
4. Linguistic Factors: Sentences
5. Organizational Factors: Plan Analysis
6. Organizational Factors: Literal Comprehension
7. Organizational Factors: Interpretive Comprehension
8. Organizational Factors: Critical Comprehension

Speaking

1. Voice Production Factors
2. Non-Verbal Factors
3. Linguistic Factors: Speech Sounds
4. Linguistic Factors: Words
5. Linguistic Factors: Sentences
6. Organizational Factors: Ideational Fluency
7. Organizational Factors: Plan Analysis
8. Organizational Factors: Organization of Details

In examining the listening domain categories, it is apparent that three were omitted from the original classification: (a)

Non-Verbal Factors; (b) Ideational Fluency Factors; and (c) Evaluation Factors. The category of Non-Verbal Factors was omitted since the listening tests were to be administered through audio-tape recordings, and these factors could therefore not be assessed.

Considerable discussion ensued concerning the domain of Ideational Fluency. While there is no doubt that the objectives in this domain were considered important by the Steering Committee and the Research Team, it was felt that these objectives were better tested in the speaking mode.

Concerning the factors involved in self-evaluation, it was felt that listeners are unlikely to be able to evaluate accurately their own listening during a communication situation. It seemed to us that the only time young listeners would likely concern themselves with self-evaluation would be when their purpose was to improve their listening skills. Since the conscious self-improvement of listening skills was not a prioritized objective, this domain was eliminated.

In examining the speaking domain categories, it is apparent that two were omitted from the original classifications: (a) Environmental Sound Factors; and (b) Evaluation Factors. The category of Environmental Sound Factors was omitted since it was considered applicable mainly to listening and should therefore be tested only in the listening mode.

As in listening, considerable discussion ensued concerning the factors involved in the Self-Evaluation domain. While self-evaluation is a prime teaching device, the conscious self-improvement of speaking skills was not prioritized highly. In addition, in order to evaluate accurately self-communication, the use of a feedback device such as audio or video tape is essential. Both time and financial constraints precluded the use of these for student self-evaluation during the testing situation. Thus the domain of

Evaluation Factors was omitted for the Speaking Tests.

Analysis of the Pilot Test data revealed that no changes were needed in the domains for the speaking pilot and field tests. However, some changes were found to be necessary in the listening test domains for the Pilot and Field Tests. First, both teacher and student comments revealed that they felt the Listening Tests at all grade levels were too long. In order to remedy this, a number of objectives were either combined with others or omitted. These omissions served another purpose as well. Following the pilot testing in listening, it was decided that at least two items should be used to test each objective. However, a test that included two items per each of 30 objectives, for example, would involve at least 60 items. This number, along with stimulus materials, would make the tests too long. Thus the omissions allowed for the shortening of the tests. Further, certain objectives were found to overlap, and could be combined. This was done where appropriate.

The data analysis also suggested that the Organizational Factors: Plan Analysis domain could be eliminated. In this case, the addition of one objective (5C1P) allowed for the elimination of objectives concerned with knowing the topic and indicating the speaker's purpose, and indicating the communication level. Thus, the elimination of these three objectives made it possible to eliminate entirely the domain of Plan Analysis.

Concerning the attitude testing, in the original domains three aspects of communication attitudes were considered important to communication effectiveness: (a) Motivation; (b) Image of self and others; and (c) Willingness to communicate. However, only two objectives were prioritized: (a) experiences a level of arousal which will facilitate communication (motivation) and (b) possesses a positive image of him/herself as a communicator. The literature review revealed that these two attitudes seem to stem from basic communication anxiety. Thus a test to measure this aspect of communication behavior was sought. A final choice of two instruments developed by McCroskey (1977) was made. These were The Personal Report of Communication Fear (PRCF) for administration at the Grade 3 and 6 levels, and The Personal Report of Communication Anxiety (PRCA) at the Grade 9 and 12 levels. These tests are self-report inventories scaled by the Thurstone method of Equal-Appearing Intervals.

ITEM DEVELOPMENT

In the Proposal to MACOSA for the development of the listening and speaking test instruments, the Research Team identified a number of considerations which research reviews suggested were of importance. These included: (a) objectives and measurement devices should be prepared in terms of oral reception of both spontaneous and prepared spoken language and written language read aloud, and in terms of the oral expression involving both spontaneous and prepared spoken language and written language read aloud; (b) objectives and measurement devices should reflect concern with the verbal aspects of listening and speaking (vocabulary, syntax, phonology, etc.) as well as with their nonverbal aspects; and (c) measurement devices should assess communication in controlled but natural situations.

In Chapter 1 of this Report, a number of recommendations were presented that were contained in two reviews of language arts literature developed for MACOSA in 1977. These included the following: (a) assessment of oral expression/reception should focus on the elements of spoken language that make it different from written language; and (b) oral expression (speaking competence) should be tested in individual or small group situations. Additional guidelines were provided in a paper on the testing of listening by Nancy Mead (1978) of the National Assessment of Educational Progress.

These included: (a) short, interesting listening stimuli should be used; (b) extraneous factors that contribute to item bias should be considered, such as language level, stimulus quality, etc.; and (c) focus should be on skills which are unique and central to listening.

In Chapter 1, some additional guidelines under which the Research Team functioned were delineated. These included: (a) the tests would be about one hour long in individual student time; (b) they should include a medley of subtests; (c) they should be criterion referenced; (d) technological aids should be used where feasible; (e) test situations should capture the performance of students in natural situations, with emphasis on both formal and semi-formal classroom situations; (f) skills should be tested directly, with students providing or responding to actual talk; (g) consideration should be given to using parts of all of available listening, speaking, and attitude tests.

The Proposal outlined the item development tasks as follows in Activity 4 (page 12), Activity 5 (page 14), and Activity 6 (page 15):

> ACTIVITY 4: By April 1, 1978, the Research Team will complete a set of Objectives/Items Matrices for each test, indicating cells for which existing items can be used as well as cells for which additional items must be prepared. The Research Team will also provide a Test Blueprint for each test.
>
> 4.1: The Research Team will prepare eight (8) Objectives/Items Matrices, one each for listening/speaking at each of four (4) grade levels: 3, 6, 9, and 12. These will contain available usable or revisable test items for measuring high priority objectives.
>
> 4.2: The Research Team will provide Test Blueprints indicating the number of items that should be included for each listening/speaking objective

at each grade level in order to achieve appropriate balances.

ACTIVITY 5: By April 30, 1978, the Research Team will construct an Item Pool which will contain potential items for the measurement of each objective as determined as a result of Activity 4.

ACTIVITY 6: By May 15, 1978, the Research Team will pilot the items in the Item Pool as well as an Administrator Checklist devised by the Team for use during the pilot testing. Upon completion of the pilot testing, the Research Team will analyze the Pilot Test and Checklist data to provide baseline information for the production of the Field Test instruments.

Thus, with the objectives selected, prioritized, and validated, the Research Team set itself to the task of developing and selecting items for the listening and speaking tests at each of Grades 3, 6, 9, and 12. Since the techniques employed and the decisions made concerning item development and selection differed somewhat for the listening and speaking tests, these are discussed separately below.

LISTENING: DEVELOPMENT OF THE PILOT TESTS

Earlier in this chapter, on page 98, a Flow Chart is provided which indicates how specific objectives can be related to any stimulus, audience, mode, purpose and communicator, at any level of stimulus and communicator sophistication. This Flow Chart served as the basis for the development of the stimulus subtests.

Language Types

First, a decision was made as to the percentage of items on each subtest that would be concerned with the three types of language that the test was designed to assess: written, spontaneous, and spontaneous planned. The Pilot Test content is provided in Table 3-1.

Table 3-1

Listening Pilot Test Items by Grade and Language Type

Grade	Language Type	Total No. of Items	Percent of Items on Test
3	Non-Language	6	11
	Written	19	52
	Spontaneous	16	30
	Planned Spontaneous	4	7
	Total (Grade 3)	55	100
6	Written	14	28
	Spontaneous	25	50
	Planned Spontaneous	11	22
	Total (Grade 6)	50	100
9	Written	21	35
	Spontaneous	23	38
	Planned Spontaneous	16	27
	Total (Grade 9)	60	100
12	Written	22	35
	Spontaneous	23	38
	Planned Spontaneous	16	27
	Total (Grade 12)	61	100

At the Grade 3 level, the non-language aspect of environmental sounds comprised 11% of the test instrument. Language meant to be read aloud comprised 52%, as story telling, reading aloud, and the like comprise a large part of early language arts programs. Since planned spontaneous language activities are minimal at the primary levels, listening to such language comprised only 7% of the test instrument.

At the Grade 6 level, the amount of listening to story-type material is generally reduced in classrooms, and there is more spontaneous talk. While there is more planned spontaneous talk (giving oral reports, etc.) at this level than at Grade 3, it is still less than at the Grade 9 and 12 levels. Thus the proportion of listening to spontaneous and planned spontaneous language in these tests was varied in terms of curriculum expectations.

Next, it was necessary to decide on the types of stimulus materials for which items could be written. In general, stimulus materials were chosen that reflected the type of listening students needed to do in classrooms--listening to prose and poetry (written language), listening to teacher talk, and listening to student talk in terms of both spontaneous (discussion and conversation) and planned spontaneous (oral reporting) language. In addition, it was decided to add some stimulus materials with which students might be concerned both in and out of the classroom, such as a typical news report which they might use as a model for their own news reporting in classes, and radio talks which might also serve as such models. Once these stimulus materials were selected, it was necessary to decide in what order to place them in the tests. This decision was arbitrary, as the Research Team had no literature or research findings to guide them. It was therefore decided to begin each test at each grade level with poetry stimulus material but to vary the other stimulus materials on an ad hoc basis. In this way the stimulus subtests were derived and presented.

Stimulus Subtests

The next step was to make a decision in terms of number of items for each stimulus subtest. The Pilot Test content in terms of subtest number, stimulus format, language type, and number and percentage of items for each stimulus subtest is provided in Table 3-2.

Table 3-2

Listening Pilot Test Content by Stimulus Subtest, Stimulus Format, Language Type, Item Numbers, and Number and Percent of Items

Grade	Stimulus Subtests	Stimulus Format	Language Type	Item Numbers	Total No. of Items	Percent of Items
3	1	Environmental Sounds	Non	1-6	6	11
	2	Poetry	Written	7-11	5	9
	3	Student Talk				
		Anecdote	Planned Spontaneous	12-15	7	13
		Recipe	Spontaneous	16-18		
	4	Teacher Talk	Spontaneous	19-25	7	13
	5	Prose: Story	Written	26-49	24	42
	6	Conversation	Spontaneous	50-55	6	11
		Total (Grade 3)			55	100
6	1	Poetry	Written	1-7	7	14
	2	News Report	Written	8-14	7	14
	3	Teacher Talk	Spontaneous	15-22	8	16
	4	Student Council Discussion	Spontaneous	23-30	8	16
	5	Student Talk Oral Report	Planned Spontaneous	31-35	5	10
	6	Radio Talk	Planned Spontaneous	36-41	6	12
	7	Conversation	Spontaneous	42-47	6	12
	8	Job Application	Spontaneous	48-50	3	6
		Total (Grade 6)			50	100
9	1	Poetry	Written	1-5	5	08
	2	News Report	Written	6-15	10	17
	3	Teacher Talk	Spontaneous	16-28	13	21
	4	Education	Planned Spontaneous	29	1	2
	5	Job Application	Spontaneous	30-32	3	5
	6	King Richard	Written	33-38	6	10
	7	Student Conversation	Spontaneous	39-45	7	12
	8	Oral Report Hypothermia	Planned Spontaneous	46-52	7	12
		Radio Talk	Planned Spontaneous	53-60	8	13
		Total (Grade 9)			60	100
12	1	Poetry	Written	1-6	6	08
	2	News Report	Written	7-16	10	17
	3	Teacher Talk	Spontaneous	17-29	13	21
	4	Education	Planned Spontaneous	30	1	2
	5	Job Application	Spontaneous	31-33	3	5
	6	King Richard	Written	34-39	6	10
	7	Student Discussion	Spontaneous	40-46	7	12
	8	Prose Never Cry Wolf	Written	47-54	8	13
	9	Student Talk	Planned Spontaneous	55-61	7	12
		Total (Grade 12)			61	100

Item Format and Selection

In order to keep the testing conditions as standardized as possible from grade to grade, it was decided to use a multiple-choice item format for all items. In order to keep the number of items manageable in terms of a one-hour test (exclusive of administration time), it was also decided to limit the number of objectives that could be tested and to attempt to use as few items as possible to test each objective. However, since these were Pilot Tests. it was clear that some items would have to be eliminated or revised. Thus a range of number of items was provided for assessing each objective. (Later, in the Field Tests the number of items per objective was limited to approximately two.)

Since no commercial or research-based listening tests were found that could be consistently used either in part or in their entirety, all of the stimulus materials were selected by the Research Team and all of the items were written by them.

In accord with Activities 4 and 5 in the Proposal, Pilot Test Blueprints were developed for each test at each grade level. These Blueprints are provided in Tables 3-3 to 3-6. Table 3-7 summarizes the Pilot Test objectives and items in the form of a matrix for all grade levels.

Administration Time

The administration time in minutes for each pilot subtest at each grade level is provided in the chart on page 127.
The times exceeding one hour were included to allow for the testing of a sufficient number of items in the event that some would need to be dropped.

Table 3-3

Listening Pilot Test Blueprint, Grade 3

Domain #	Domain	Obj. #	# Items Per Obj.	# Obj. in Dom.	% Obj. in Dom.	# Items in Dom.	% Items in Dom.
1	Environmental Sound Factors	1A1	3	3	14	6	11
		1A2	2				
		1B3	1				
2	Voice Production Factors	2B1	6	3	14	12	22
		2D1	4				
		2E1	2				
3	Linguistic Factors - Words	4B1	4	3	14	7	13
		4B3	1				
		4B7	2				
4	Linguistic Factors - Sentences	0	0	0	0	0	0
5	Organizational Factors - Plan Analyses	5B1	1	1	05	1	02
6	Organizational Factors - Literal Comprehension	5C1A	3	4	19	11	20
		5C1C	1				
		5C1E	5				
		5C1H	2				
7	Organizational Factors - Interpretational Comprehension	5C2C	4	4	19	13	23
		5C2D	3				
		5C2F	3				
		5C2H	3				
8	Organizational Factors - Critical Comprehension	5C3A	2	3	14	5	9
		5C3D	2				
		5C3M	1				
	Totals		55	21	99	55	100

Table 3-4

Listening Pilot Test Blueprint, Grade 6

Domain #	Domain	Obj. #	# Items Per Obj.	# Obj. in Dom.	% Obj. in Dom.	# Items in Dom.	% Items in Dom.
2	Voice Production Factors	2B1 2D1 2E1	5 5 2	3	15	12	24
3	Linguistic Factors - Words	4B1 4B3 4B7	4 3 2	3	15	9	18
4	Linguistic Factors - Sentences	0	0	0	0	0	0
5	Organizational Factors - Plan Analyses	5B1 5B2 5B5	1 2 3	3	15	6	12
6	Organizational Factors - Literal Comprehension	5C1A 5C1C 5C1E 5C1H	2 3 2 2	4	20	9	18
7	Organizational Factors - Interpretational Comprehension	5C2C 5C2D 5C2F 5C2H	2 3 2 2	4	20	9	18
8	Organizational Factors - Critical Comprehension	5C3A 5C3D 5C3M	1 2 2	3	15	5	10
	Totals		50	20	100	50	100

Table 3-5

Listening Pilot Test Blueprint, Grade 9

Domain #	Domain	Obj. #	# Items Per Obj.	# Obj. in Dom.	% Obj. in Dom	# Items in Dom.	% Items in Dom.
2	Voice Production Factors	2B1	2	3	10	4	7
		2D1	1				
		2E1	1				
3	Linguistic Factors - Words	4B1	4	3	10	10	17
		4B3	3				
		4B7	3				
4	Linguistic Factors - Sentences	4C2	2	3	10	6	10
		4C4	2				
		4C5	2				
5	Organizational Factors - Plan Analyses	5B1	1	3	10	7	11
		5B2	2				
		5B5	4				
6	Organizational Factors - Literal Comprehension	5C1A	2	6	20	13	22
		5C1C	2				
		5C1D	2				
		5C1E	2				
		5C1G	3				
		5C1H	2				
7	Organizational Factors - Interpretational Comprehension	5C2C	2	6	20	14	23
		5C2D	3				
		5C2F	2				
		5C2H	2				
		5C2K	2				
		5C2O	3				
8	Organizational Factors - Critical Comprehension	5C3A	1	6	20	6	10
		5C3B	1				
		5C3E	1				
		5C3H	1				
		5C3J	1				
		5C3O	1				
	Totals		60	30	100	60	100

Table 3-6

Listening Pilot Test Blueprint, Grade 12

Domain #	Domain	Obj. #	# Items Per Obj.	# Obj. in Dom.	% Obj. in Dom.	# Items in Dom.	% Items in Dom.
2	Voice Production Factors	2B1 2D1 2E1	2 1 1	3	10	4	07
3	Linguistic Factors - Words	4B1 4B3 4B7	3 3 5	3	10	11	18
4	Linguistic Factors - Sentences	4C2 4C4 4C5	2 2 1	3	10	5	08
5	Organizational Factors - Plan Analysis	5B1 5B2 5B5	1 2 5	3	10	7	11
6	Organizational Factors - Literal Comprehension	5C1A 5C1C 5C1D 5C1E 5C1G 5C1H	2 2 4 1 2 2	6	20	13	21
7	Organizational Factors - Interpretational Comprehension	5C2C 5C2D 5C2F 5C2H 5C2K 5C2O	3 4 2 1 2 3	6	20	15	25
8	Organizational Factors - Critical Comprehension	5C3A 5C3B 5C3E 5C3H 5C3J 5C3O	1 1 1 1 1 1	6	20	6	10
	Totals		61	30	100	61	100

Table 3-7

Listening Pilot Test Objectives/Items Matrices
Grades 3, 6, 9, 12

Domain #	Obj. #	# of Items		Grade Level # of Items		# of Items		# of Items	
		3	DT	6	DT	9	DT	12	DT
1. Environmental Sound Factors N=3	1A1	3		0		0		0	
	1A2	2		0		0		0	
	1B3	1	6	0	0	0	0	0	0
2. Voice Production Factors N=3	2B1	6		5		2		2	
	2D1	4		5		1		1	
	2E1	2	12	2	12	1	4	1	4
3. Linguistic Factors - Words N=3	4B1	4		4		4		3	
	4B3	1		3		3		3	
	4B7	2	7	2	9	3	10	5	11
4. Linguistic Factors - Sentences N=3	4C2	0		0		2		2	
	4C4	0		0		2		2	
	4C5	0	0	0	0	2	6	1	5
5. Organizational Factors - Plan Analyses N=3	5B1	1		1		1		1	
	5B2	0		2		2		2	
	5B5	0	1	3	6	4	7	4	7
6. Organizational Factors - Literal Comprehension N=6	5C1A	3		2		2		2	
	5C1C	1		3		2		2	
	5C1D	0		0		2		4	
	5C1E	5		2		2		1	
	5C1G	0		0		3		2	
	5C1H	2	11	2	9	2	13	2	13
7. Organizational Factors - Interpretative Comprehension N=6	5C2C	4		2		2		3	
	5C2D	3		3		3		3	
	5C2F	3		2		2		2	
	5C2H	3		2		2		1	
	5C2K	0		0		2		2	
	5C2O	0	13	0	9	3	14	3	15
8. Organizational Factors - Critical Comprehension N=7	5C3A	2		1		1		1	
	5C3B	0		0		1		1	
	5C3D	2		2		0		0	
	5C3E	0		0		1		1	
	5C3H	0		0		1		1	
	5C3J	0		0		1		1	
	5C3M	1	5	2	5	1	6	1	6
Totals	34	55		50		60		61	

Grade 3 = 21 Objectives
Grade 6 = 20 Objectives
Grade 9 = 30 Objectives
Grade 12 = 30 Objectives

Grade	Preparation (Minutes)	Test Time (Minutes)	Break (Minutes)	Total Time (Minutes)
3	15	55	5	75
6	15	55	5	75
9	15	75	5	95
12	15	75	5	95

Administration Manual (Teacher's Manual)

A Teacher's Manual was developed for each Listening Test at each grade level. Each Manual contained introductory material which provided: (a) a simple explanation of the purpose, format, and content of the tests; (b) administration instructions other than those provided on the tape; and (c) a tape script of the entire test. Answer keys were not provided, as all test administration was done by certified substitute teachers; marking was done by computer; and testing material was collected by the test administrators and returned to the Research Team immediately after the tests were given.

Stimulus Materials

Tables 3-8 to 3-11 indicate the stimulus materials for each Pilot Test at each grade level, the source of the material, and the speaker/reader.

Student Test Booklets

In accordance with typical practice in the testing of listening, the Student Test Booklets for each test at each grade level

Table 3-8

Listening Pilot Test Scripts and Sources, Grade 3

Subtest	Content	Source	Language
1. Environmental Sounds	Sounds	Originals and Sound Bands	Non-Language
2. Poetry	John Ciardi, "The Man in the Onion Bed." Stanza 1 and 2.	Wm. Cole (ed.), Oh, What Nonsense. N.Y: Viking Press, 1966. 625 Madison Ave. New York 0022 (Macmillan Company, 70 Bond Street, Toronto M5B 1X3).	Written
3a. Student Talk: Anecdote	"The Day the Bear Came to Lunch.	E. Plattor et. al. English Skills Program, Bk. 1. Toronto: Gage Educational Publishing Ltd., 1977, p. 79. (Adapted)	Prepared Spontaneous
3b. Student Talk: Recipe	"Spufoni"	Jane Martel (Ed.) Smashed Potatoes. Boston: Houghton-Mifflin, 1974. (Adapted)	Spontaneous
4. Teacher Talk	Field Trip to Roundup '78	Original Piece	Spontaneous
5. Prose: Narrative	Grace Lurton, "The Cuckoo Clock."	Child Life, June/July, 1976, pp. 14-17 Review Publishing Comp. 1100 Waterway Blvd. Indianapolis, Ind. 46202	Written
6. Conversation	Pieces of Children's Conversations with their Parents	Original pieces	Spontaneous

Table 3-9

Listening Pilot Test Scripts and Sources, Grade 6

Subtest	Content	Source	Language
1. Poetry	Joan Raeside, "The Rich Spider."	Canadian Children's Magazine #2, Autumn, 1976, p. 40.	Written
2. News Report	"Garage Haunting by a Poltergeist"	Ghost Special (no author, publisher, or date given), page 19, Adapted.	Written
3. Teacher Talk	Gymnastics Event	Original Piece	Spontaneous
4. Job Application	Job Application	E. Plattor et. al. English Skills Prog. 1, Toronto: Gage Educational Pub. Ltd., 1977, p. 15.	Spontaneous
5. Oral Report	RCMP Police Service Dog	E. Plattor et. al., English Skills Prog. 1, Toronto: Gage Educational Pub. Ltd., 1977, p. 73.	Prepared Spontaneous
6. Student Discussion	Student Council - re: Purchase of Stereo	Original Piece	Spontaneous
7. Radio Talk	Newspaper Carrier Boys	CBC	Prepared Spontaneous

Table 3-10

Listening Pilot Test Scripts and Sources, Grade 9

Subtest	Content	Source	Language
1. Poetry	Patricia Hubbell, "Concrete Mixers."	Nancy Larrick, Piping Down the Valleys Wild; Poetry for the Young of All Ages. Delacorte Press, 1968. Dial Press - 1 Dag Hammarskjold Plaza, N.Y., 242 & 47 Street, N.Y. 10017	Written
2. News Report	Public Meeting on Youth Hostels	J. Bell & M. Roe, Action English 4. Toronto: Gage Ed. Publ. Ltd., 1973, p. 92.	Written
3. Teacher Talk	Assignment of movie review - discussion of film: Walkabout	Cy Groves - original talk	Unprepared Spontaneous
4. Job Application	Job Application	E. Plattor et. al., English Skills Program 1, Toronto: Gage Ed. Publ. Ltd., p. 15.	Unprepared Spontaneous
5a. Education Talk	Getting an Education	Original piece by Plattor & Unruh	Prepared Spontaneous
5b. Oral Report	Report on Hypothermia	E. Plattor et. al., English Skills Program 3, Toronto: Gage Ed. Publ. Ltd., pp.30, 46-49.	Prepared Spontaneous
6. Play	Shakespeare: Richard II		Written
7. Student Discussion	Discussion about a line in a poem.	"Missing Dates," Wm. Empson. Penguine Book of English Verse. John Hayward (ed) Harmonsworth: Penguin Books, 625 Madison Ave. 10022, 1956, p. 452.	Unprepared Spontaneous
8. Radio Talk	Newspaper Carrier Boys	CBC, Calgary - Jack Peach	Prepared Spontaneous

Table 3-11

Listening Pilot Test Scripts and Sources, Grade 12

Subtest	Content	Source	Language
1. Poetry	W. H. Auden, "The Unknown Citizen."	David Aloian (ed)., Poems to Read Aloud. Webster Division, McGraw Hill Book Co., McGraw Hill Ryerson Ltd., 330 Progress Avenue, Scarborough, Ontario, M1P 2Z5, 1965.	Written
2. News Report	Public Meeting on Youth Hostels	J. Bell & M. Roe, Action English 4. Toronto: Gage Ed. Publ. Ltd., 1973, p. 92.	Written
3. Teacher Talk	Assignment of movie review--discussion of film: Walkabout	Cy Groves - Original Talk	Unprepared Spontaneous
4. Job Application	Job Application	E. Plattor et. al., English Skills Program 1. Toronto: Gage Ed. Publ. Ltd., p. 15.	Unprepared Spontaneous
5a. Education Talk	Getting an Education	Original piece by Plattor & Unruh	Prepared Spontaneous
5b. Student Talk	Drugs	K. Weber. Prose or Relevance. Toronto: Methuen, 1971, pp. 59-60 (Adapted).	Prepared Spontaneous
6. Play	Shakespeare: Richard II		Written
7. Student Discussion	Discussion about a line in a poem.	"Missing Dates," Wm. Empson. Penguin Books of English Verse. John Hayward (ed) Harmonsworth: Penguin Books, 1956, p. 452.	Unprepared Spontaneous
8. Student Talk	Graduation Address	Original Talk	Prepared Spontaneous
9. Prose	Farley Mowat, Never Cry Wolf	F. Mowat, Never Cry Wolf, p. 91, McClelland & Stewart, 25 Hollinger Rd., Toronto, M4B 3G2 2330 Midland Avenue Agincourt, M1C 1D7	Written

contained only the multiple choice options for each question. In addition, the Booklets contained introductory comments and instructions about the entire test as well as subtest directions which the students could read as they listened to them on the tape recordings.

Audio Tape Recordings

All of the audio tape recordings were made in the sound studios of the University of Calgary's Faculty of Education. The stimulus materials that were not recordings made in classrooms or from radio tapes were read by in-service teachers, as were the questions and the multiple choice items. A six-second "beep" was provided after each question to key student responses and before each new question to prepare the students for the next question. The master audio tapes were produced on 1/4 x 1800' x 7' reels; speed, 7 1/2" per second, orb 19cm. Reel-to-reel copies of each master tape were made for use in the test administration.

LISTENING: DEVELOPMENT OF THE FIELD TESTS

The data analyses of the Pilot Test results indicated the need for certain minimal changes in the preparation of the Field Tests. These revisions as well as the results of the data analyses of both the pilot and field testing are described in Chapter 4. Final test revisions based on the field test data analyses are also described in Chapter 4.

Language Types

The Field Test content of the Grade 3, 6, 9, and 12 listening tests in terms of language is provided in Table 3-12.

At the Grade 3 level, the domain of Environmental Sounds was eliminated from the field test. The decision to eliminate this domain was based on pilot test data analyses which indicated that

Table 3-12

Listening Field Test Items by Grade and Language Type

Grade	Language Type	Total No. of Items	Percent of Items on Test
3	Written	18	50
	Spontaneous	14	41
	Planned Spontaneous	3	8
	Total (Grade 3)	35	99
6	Written	11	27.5
	Spontaneous	19	47.5
	Planned Spontaneous	10	25
	Total (Grade 6)	40	100.0
9	Written	17	34
	Spontaneous	19	38
	Planned Spontaneous	14	28
	Total (Grade 9)	50	100
12	Written	22	44
	Spontaneous	16	32
	Planned Spontaneous	12	24
	Total (Grade 12)	50	100

the students found the content and items very easy, thus suggesting that students at the Grade 3 level are able to comprehend environmental sounds without difficulty. Thus the percentage of language types differ at the Grade 3 level from the Pilot to the Field Test, with the largest increase in spontaneous language use. At the Grade 3 level, 30 percent of the items on the Pilot Test dealt with spontaneous language, while 41 percent of the Field Test items dealt with spontaneous language. Thus it can be seen that stimulus materials and items which dealt with environmental sounds in the Pilot Test were replaced with stimulus materials which were concerned with spontaneous language. At the other grade levels there were few changes in language type percentages between the Pilot and Field Tests.

Stimulus Subtests

The Field Test content in terms of stimulus subtest number, stimulus format, language type, and number and percentage of items for each stimulus format is provided in Table 3-13.

A comparison of the Field Test and Pilot Test charts (see pages 120 and 135) reveals little change in the number and order of the stimulus subtests. However, with the elimination of the domain of Environmental Sounds at Grade 3, there were five stimulus subtests in the Field Test as opposed to six stimulus subtests in the Pilot Test. At the Grade 6, 9, and 12 levels, changes in the order of the subtests were made to permit the closer placement of certain stimuli of the same language type. For example, at the Grade 9 and 12 levels, the Pilot stimulus subtest called "Education" was included with the "Student Talk" subtest and identified as "Oral Report 1." At the Grade 6 level, the spontaneous conversation subtest was dropped, as this was a repetition of a Grade 3 subtest which had proved sufficiently easy for the Grade 3 students to warrant its elimination as a Grade 6 subtest. At the Grade 9 level, a new piece of student discussion was added to the "Student Talk: Discussion" subtest. At the Grade 12 level, a new subtest concerned with planned spontaneous talk ("Graduation Address") was added.

Thus, at the Grade 3 level there were six stimulus Pilot subtests and five stimulus Field subtests; at the Grade 6 level there were eight stimulus Pilot subtests and seven stimulus Field subtests; at the Grade 9 level there were nine stimulus Pilot subtests and

Table 3-13

Listening Field Test Content by Stimulus Subtest, Stimulus Format, Language Type, Item Number and Number and Percent of Items

Grade	Stimulus Subtests	Stimulus Format	Language Type	Item Numbers	Total No. of Items	Percent of Items
3	1	Poetry	Written	1-4	4	11
	2	Student Talk Anecdote	Planned Spontaneous	5-7	3	8
		Recipe	Spontaneous	8-10	3	8
	3	Teacher Talk	Spontaneous	11-15	5	14
	4	Story	Written	16-29	14	41
	5	Conversation	Spontaneous	30-35	6	18
		Total (Grade 3)			55	100
6	1	Poetry	Written	1-5	5	12.5
	2	News Report	Written	6-11	6	15
	3	Teacher Talk	Spontaneous	12-16	5	12.5
	4	Student Talk Job Application	Spontaneous	17-20	4	10
	5	Student Talk Oral Report	Planned Spontaneous	21-26	6	15
	6	Student Talk Students' Council Discussion	Spontaneous	27-36	10	25
	7	Radio Talk	Planned Spontaneous	37-40	4	10
		Total (Grade 6)			40	100.0
Grade	Stimulus Subtests	Stimulus Format	Language Type	Item Numbers	Total No. of Items	Percent of Items
9	1	Poetry	Written	1-4	4	08
	2	News Report	Written	5-13	9	18
	3	Teacher Talk	Spontaneous	14-22	9	18
	4	Student Talk Job Application	Spontaneous	23-25	3	06
	5	Student Talk Oral Report 1	Planned Spontaneous	26	1	
		Oral Report 2	Planned Spontaneous	27-33	7] 8	16
	6	Play	Written	34-37	4	8
	7	Student Talk Discussion 1	Spontaneous	38-41	4	
		Discussion 2	Spontaneous	42-44	3] 7	14
	8	Radio Talk	Planned Spontaneous	45-50	6	12
		Total (Grade 9)			50	100
12	1	Poetry	Written	1-4	4	08
	2	News Report	Written	5-13	9	18
	3	Teacher Talk	Spontaneous	14-22	9	18
	4	Student Talk Job Application	Spontaneous	23-25	3	06
	5	Student Talk Oral Report 1	Planned Spontaneous	26	1	16
		Oral Report 2	Planned Spontaneous	27-33	7] 8	
	6	Play	Written	34-37	4	08
	7	Student Talk Discussion	Spontaneous	38-41	4	08
	8	Student Talk Graduation Address	Planned Spontaneous	42-45	4	08
	9	Prose	Written	45-50	5	10
		Total (Grade 12)			50	100

eight stimulus Field subtests; and at the Grade 12 level there were nine stimulus Pilot subtests and nine stimulus Field subtests.

Item Format and Selection

Tables 3-14 to 3-18 present the Listening Field Test Blueprints and Objective/Items Matrices as revised for field testing. Revisions were made in item number, item content, and number of objectives. The revisions in item number were made in order to decrease the amount of time for each test. This was necessary because of teacher and student comments as well as student behavior during the test. A comparison of these charts with the Pilot Test charts (see pages 122 to 126) reveals that, at the Grade 3 level there were 55 items on the Pilot Test and 35 items on the Field Test; at the Grade 6 level there were 50 items on the Pilot Test and 40 items on the Field Test; at the Grade 9 level there were 60 items on the Pilot test and 50 items on the Field test; and at the Grade 12 level there were 61 items on the Pilot Test and 50 items on the Field Test.

Administration Time

The Administration Time in minutes for each Field Test at each grade level is shown in the chart below.

Grade	Preparation (Minutes)	Test Time (Minutes	Break (Minutes)	Total Time (Minutes)
3	15	40	5	60
6	10	45	5	65
9	10	60	0	70
12	10	60	0	70

Table 3-14

Listening Field Test Blueprint, Grade 3

Domain #	Domain	Obj. #	# Items per Obj.	# Obj. in Domain	% Obj. in Domain	# Items in Domain	% Items in Domain
1	Voice Production Factors	2B1 2D1 2E1	2 5 2	3	21	9	25
2	Linguistic Factors: Words	4B1 4B7	3 2	2	14	5	15
3	Linguistic Factors: Sentences	0	0	0	0	0	0
4	Organizational Factors: Literal Comprehension	5C1A 5C1C 5C1E 5C1H	2 2 3 2	4	30	9	25
5	Organizational Factors: Interpretive Comprehension	5C2C 5C2F 5C2H	2 3 2	3	21	7	20
6	Organizational Factors: Critical Comprehension	5C3D 5C3O	2 3	2	14	5	15
	TOTALS	14	35	14	100	35	100

Table 3-15

Listening Field Test Blueprint, Grade 6

Domain #	Domain	Obj. #	# Items per Obj.	# Obj. in Domain	% Obj. in Domain	# Items in Domain	% Items in Domain
1	Voice Production Factors	2B1 2D1 2E1	4 2 2	3	19	8	20
2	Linguistic Factors: Words	4B1 4B3 4B7	3 2 2	3	19	7	17
3	Linguistic Factors: Sentences	0	0	0	0	0	0
4	Organizational Factors: Literal Comprehension	5C1A 5C1C 5C1E 5C1H 5C1P	2 2 3 2 3	5	30	12	30
5	Organizational Factors: Interpretive Comprehension	5C2C 5C2F 5C2H	2 3 2	3	19	7	17
6	Organizational Factors: Critical Comprehension	5C3D 5C30	2 4	2	12	6	15
	TOTALS	16	40	16	99	40	100

Table 3-16

Listening Field Test Blueprint, Grade 9

Domain #	Domain	Obj. #	# Items per Obj.	# Obj. in Domain	% Obj. in Domain	# Items in Domain	% Items in Domain
1	Voice Production Factors	2B1 2D1	2 3	2	10	5	10
2	Linguistic Factors: Words	4B1 4B3 4B7	3 2 3	3	15	8	16
3	Linguistic Factors: Sentences	4C2 4C5	2 2	2	10	4	8
4	Organizational Factors: Literal Comprehension	5C1A 5C1C 5C1D 5C1E 5C1G 5C1H 5C1P	3 3 2 2 4 2 2	7	35	18	36
5	Organizational Factors: Interpretive Comprehension	5C2C 5C2F 5C2H	3 2 2	3	15	7	14
6	Organizational Factors: Critical Comprehension	5C3B 5C3E 5C30	2 2 4	3	15	8	16
	TOTALS	20	35	20	100	50	100

Table 3-17

Listening Field Test Blueprint, Grade 12

Domain #	Domain	Obj. #	# Items per Obj.	# Obj. in Domain	% Obj. in Domain	# Items in Domain	% Items in Domain
1	Voice Production Factors	2B1 2D1	2 2	2	10	4	08
2	Linguistic Factors: Words	4B1 4B3 4B7	2 2 3	3	15	7	14
3	Linguistic Factors: Sentences	4C2 4C5	2 2	2	10	4	08
4	Organizational Factors: Literal Comprehension	5C1A 5C1C 5C1D 5C1E 5C1G 5C1H 5C1P	2 2 3 2 3 1 4	7	35	17	34
5	Organizational Factors: Interpretive Comprehension	5C2C 5C2F 5C2H	3 3 3	3	15	9	18
6	Organizational Factors: Critical Comprehension	5C3B 5C3E 5C3O	2 3 4	3	15	9	18
	TOTALS	20	50	20	100	50	100

Table 3-18

Listening Field Test Objectives/Items Matrices,
Grades 3, 6, 9, 12

		Grade Level							
Domain #	Objective #	3	DT	6	DT	9	DT	12	DT
1. Voice Production Factors N=3	2B1	2		4		2		2	
	2D1	5		2		3		2	
	2E1	2	9	2	8	0	5	0	4
2. Linguistic Factors - Words N=3	4B1	3		3		3		2	
	4B3	0		2		2		2	
	4B7	2	5	2	7	3	8	3	7
3. Linguistic Factors - Sentences N=2	4C2	0		0		2		2	
	4C5	0	0	0	0	2	4	2	4
4. Organizational Factors - Literal Comprehension N=7	5C1A	2		2		3		2	
	5C1C	2		2		3		2	
	5C1D	0		0		2		3	
	5C1E	3		3		2		2	
	5C1G	0		0		4		3	
	5C1H	2		2		2		1	
	5C1P	0	9	3	12	2	18	4	17
5 Organizational Factors - Interpretive Comprehension N=3	5C2C	2		2		3		3	
	5C2F	3		3		2		3	
	5C2H	2	7	2	7	2	7	3	9
6. Organizational Factors - Critical Comprehension N=4	5C3B	0		0		2		2	
	5C3D	2		2		2		3	
	5C3E	0		0		4		4	
	5C3O	3	5	4	6	0	8	0	9
Totals	22	35		40		50		50	

Grade 3 = 14 objectives

Grade 6 = 16 objectives

Grade 9 = 20 objectives

Grade 12 = 20 objectives

A comparison of the Pilot and Field Test administration times reveals that the testing period was shortened at each grade level. (See page 127) for Pilot Test administration times.

At the Grade 3 level, the actual Field Test time was cut by 15 minutes, while at Grade 6 the time was cut by 10 minutes. The preparation time, which included distributing materials, filling out demographic information on answer sheets, etc., was cut by about five minutes in the field testing. At the Grade 9 and 12 levels, the preparation time was cut by about five minutes as well, and the actual testing time was cut by 10 minutes at Grade 9 and 15 minutes at Grade 12.

It was also found during the pilot testing that the break time was dispensible at Grades 6-9, thus cutting the total testing time at these levels by another five minutes.

Administration Manual (Teacher's Manual)

Few changes were made in the Field Test Administration Manuals. The only revisions involved changes in wording for purposes of clarity.

Stimulus Materials

Tables 3-19 to 3-22 list the stimulus materials for each field subtest at each grade level, along with the source of the material.

Student Test Booklets

As in the pilot testing, the Student Test Booklets for each test at each grade level contained only the multiple choice options for each question. However, a basic format change was made in the Field Test.

In the pilot Student Test Booklets, the directions for each stimulus subtest were followed immediately by the multiple choice

Table 3-19

Listening Field Test Scripts and Sources, Grade 3

Subtest	Content	Source
1. Poetry	John Ciardi, "The Man in the Onion Bed." Stanza 1 & 2.	Wm. Cole (ed.), Oh, What Nonsense. N.Y.: Viking Press, 1966. 625 Madison Ave. New York 0022 (Macmillan Comp., 70 Bond Street, Toronto M5B 1X3).
2. Student Talk: Anecdote	"The Day the Bear Came to Lunch."	E. Plattor et. al. English Skills Program, Bk. 1. Toronto: Gage Educational Publishing Ltd., 1977, p. 79. (Adapted)
2. Student Talk: Recipe	"Spufoni"	Jane Martel (Ed.) Smashed Potatoes. Boston: Houghton-Mifflin, 1974. (Adapted)
4. Prose: Narrative (Story)	Grace Lurton, "The Cuckoo Clock."	Child Life, June/July, 1976 pp. 14 - 17. Review Publishing Comp. 1100 Waterway Blvd. Indianapolis, Ind. 46202

Table 3-20

Listening Field Test Scripts and Sources, Grade 6

Subtest	Content	Source
1. Poetry	Joan Raeside, "The Rich Spider."	Canadian Children's Magazine #2, Autumn, 1976, p. 40.
2. News Report	"Garage Haunting by a Poltergeist."	Ghost Special (no author, publisher, or date given), page 19, Adapted.
4. Job Application	Job Application	E. Plattor et al. English Skills Program 1, Toronto: Gage Educational Pub. Ltd., 1977, p. 15.
5. Oral Report	RCMP Police Service Dog	E. Plattor et al., English Skills Program 1, Toronto: Gage Educational Pub. Ltd., 1977, p. 73
7. Radio Talk	Newspaper Carrier Boys	CBC, Calgary - Jack Peach.

Table 3-21

Listening Field Test Scripts and Sources, Grade 9

Subtest	Content	Source
1. Poetry	Patricia Hubbell, "Concrete Mixers."	Nancy Larrick, Piping Down the Valleys Wild; Poetry for the Young of All Ages. Delacorte Press, 1968. Dial Press - 1 Dag Hammarskjold Plaza N.Y., 242 & 47 Street, N.Y. 10017
2. News Report	Public Meeting on Youth Hostels	J. Bell & M. Roe, Action English 4. Toronto: Gage Ed. Pub. Ltd., 1973. p. 92
4. Job Applications	Job Application	E. Plattor et. al., English Skills Program 1. Toronto: Gage Ed. Pub. Ltd., p. 15.
5. Oral Report	Report on Hypothermia	E. Plattor et. al., English Skills Program 3. Toronto: Gage Ed. Pub. Ltd. pp. 30, 46 - 49.
6. Play	Shakespeare: Richard II	
7. Student discussion	Discussion about a line in a poem.	"Missing Dates", Wm. Empson. Penguin Book of Eng. Verse. John Hayward (ed) Harmonsworth: Penguin Books, 625 Madison Ave. 10022 1956, p. 452.
8. Radio Talk	Newspaper Carrier Boys	CBC, Calgary - Jack Peach

Table 3-22

Listening Field Test Scripts and Sources, Grade 12

Subtest	Content	Source
1. Poetry	W. H. Auden, "The Unknown Citizen."	David Aloian (ed.), Poems to Read Aloud. Webster Division, McGraw Hill Book Co., McGraw Hill Ryerson Ltd., 330 Progress Avenue, Scarborough, Ontario, M1P 2Z5, 1965.
2. News Report	Public Meeting on Youth Hostels	J. Bell & M. Roe, Action English 4. Toronto: Gage Ed. Pub. Ltd., 19731, p. 92.
4. Job Applications	Job Application	E. Plattor et. al., English Skills Program 1. Toronto: Gage Ed. Pub. Ltd., p. 15.
6. Play	Shakespeare: Richard II	
7. Student discussion	Discussion about a line in a poem.	"Missing Dates", Wm. Empson. Penguin Book of Eng. Verse. John Hayward (ed) Harmonsworth: Penguin Books, 1956, p. 452.
7. Student Talk	Drugs	K. Weber. Prose or Relevance. Toronto: Methuen, 1971, pp. 59-60 (Adapted).
8. Prose	Farley Mowat, Never Cry Wolf	F. Mowat, Never Cry Wolf, p. 91, McClelland & Stewart 25 Hollinger Road, Toronto M4B 3G2 2330 Midland Avenue, Agincourt M1S 1P7

options for each of the subtest questions. As teachers and test administrators observed the students listening to these directions, they found the students' listening distracted since they were reading the multiple choice options for the questions which were to follow. The teachers and administrators therefore suggested that the directions for each subtest be printed on pages separate from those which carried the subtest options. This was done in the Field Test Student Test Booklets and worked quite well.

Audio Tape Recordings

As with the pilot testing, all of the audio tape recordings were made in the sound studios of the University of Calgary's Faculty of Education. The stimulus materials that were not recorded in regular classrooms or from radio talks were read by in-service teachers, as were the questions and multiple choice options. A six-second "beep" was provided after each question to key student responses and before each new question to prepare the students for the next question. The master audio tapes were produced on 1/4 x 1800' x 7' reels; speed, 7 1/2" per second; orb 19cm. The master tapes were reproduced on cassettes for use in the test administration.

SPEAKING: DEVELOPMENT OF THE PILOT TESTS

When development of the Pilot Speaking Tests was undertaken, consideration was given to the information contained in the Flow Chart (Figure 1 , page 98). Additional criteria for the selection of objectives and items for the Pilot Speaking Tests were:

1. Do the objective and items reflect the functions of speech considered important to communication in the broad sense? (Brown and Van Riper's (1966) analysis of the critical functions (or tasks) of human communication was accepted as a definition of these functions.)
2. Are behaviors to be tested those which could reasonably be expected of students? (Since communication training has not been an established part of the curriculum in Alberta this fact had to be taken into consideration.)

The use of the Flow Chart in Figure 3-1 served to ensure the inclusion of those elements of situation, purpose, audience, etc. considered to be important to the process-relationship nature of speaking.

It was necessary to limit the range of communication events which could be assessed within the time limits of the test. The area of group or interpersonal communication, although considered of vital importance, thus had to be omitted.

Testing of both knowledges and skills was considered. Prior to Pilot Testing, separate sets of objectives were developed for both knowledges and skills. These were later combined into a single set of knowledge and skill objectives for pilot and field testing. In general, the knowledge objectives were tested in written form while skill objectives were tested by use of rating scales administered by experienced raters.

Two main subtests, Speaking (Written) and Speaking (Oral) were developed.

Speaking (Written)

The Speaking (Written) Tests at all levels attempted to assess student knowledges of how voice production, nonverbal, and linguistic factors could be used to convey the meaning, mood and interest of a written selection when read aloud to an audience.

At the Grade 9 and 12 levels the test also included items designed to assess the student's knowledge of speech organization. These items were derived from a scrambled outline stimulus. Literature reviewed suggested that one of the most important elements in extemporaneous speaking is the ability to order materials. This includes the development of ideas, and the organizing of proofs in terms of audience requirements, time constraints, and selection of language. Many of these concepts are included in the language arts/English curriculum, so a transfer to oral speech construction could reasonably be expected.

The Speaking (Written) Tests at all levels included objectives from each domain. A multiple-choice format was selected for these tests and items were derived from stimulus materials discussed below.

Speaking (Oral)

At the Grade 3 level the domain of Linguistic Factors/Sounds was tested. The Articulation Test, prepared to assess this factor, was based on research findings in the area of speech sound motivation and misarticulations common to elementary school children.

Spontaneous speech was tested at all grade levels. The stimulus for spontaneous speech at the Grade 3 and 6 levels consisted of a picture sequence constructed by the Research Team to which the student was required to respond by telling a story about the events depicted. The only structure imposed on the task was that the story have a beginning, a middle and an ending. This procedure for testing spontaneous oral language has been used by a number of investigators.

To evaluate student oral performance, rating scales were produced which included items keyed to objectives in each domain. This procedure was used at both Grade 3 and 6 levels. Picture sequences developed by the Research Team for both grades are included in the Student Test Booklets.

At the Grade 9 and 12 levels spontaneous and spontaneous-prepared speech were integrated into one task. This task consisted of the preparation and presentation of an extemporaneous speech, defined as such since preparation time was allowed. Spontaneity lay not so much in the organization of ideas as in language usage. In order to keep the stimulus constant, it was necessary to prescribe a topic for discussion. The topic of "Hiking" was chosen as having sufficient universality and interest. Students were given 45 minutes to prepare. A number of supporting materials in the form of statistics, quotations, etc. were provided in the Test Booklet. Students were asked to consider the factors of language levels, situation and audience as well as the classic divisions of informing, persuading and entertaining as outlined in the Flow Chart. The speech was limited to three minutes and was to be presented to an audience of peers in the presence of a tester. The tester would evaluate the speech as it was presented and also tape-record it for later analysis. To evaluate this task, rating scales were produced which included items keyed to objectives in each domain.

The testing of oral reading at all grade levels was also undertaken. For the purpose of this test the basic requirements of oral reading--the conveying of the meaning and the mood of a selection of literature to an audience--were considered. Consideration was given to testing oral modes with which the student might have had previous experience. It was assumed that most students would have had experience in the oral reading of poetry or prose in the classroom.

Both prose and poetry selections were selected for testing at the Grade 3 and 6 levels. Poetry, as demanding more sophicticated reading skills, was the mode tested at the Grade 9 and 12 levels. Prose and/or poetry selections suited to each grade level were selected. The same poems or prose selections were used in both the written and the oral tests. This use of the same materials for both tests had the advantage of saving preparation time and of relating objectives in the linguistic domain to preparation for oral reading. Authorities in the field agree that, since comprehension is necessary to expression, it is impossible to evaluate voice production factors such as intonation, pause, rate, etc. in sight-reading. Items which tested the knowledge of denotative and connotative meanings, stylistic devices, sequencing, etc. on the written test were seen to prepare the student for the task of conveying the meaning and mood of the selection to the audience.

Rating scales were provided to assess oral reading at all grade levels.

Construction of Rating Scales for Spontaneous Speech and Oral Reading. Behavioral objectives which describe measurable, observable student behaviors must provide meaningful and precise criteria about the performance level of students. Such criteria were sought by examining a number of rating scales used in the evaluation of both oral reading and speaking. Items which were seen to measure the stated objectives of the MACOSA test were selected and other items constructed.

Table 3-23 summarizes the Pilot Test items by grade and language type.

Table 3-23
Speaking Pilot Test Items by Grade
and Language Type

Grade	Language Type	Total No. of Items	Percentage of Items on Test
3	Written	10	15
	Spontaneous	36	51
	Oral Reading	24	34
	Total	70	100
6	Written	15	20
	Spontaneous	38	50
	Oral Reading	23	30
	Total	76	100
9	Written	25	27
	Spontaneous	55	58
	Oral Reading	14	15
	Total	94	100
12	Written	19	21
	Spontaneous	56	63
	Oral Reading	14	16
	Total	89	100

Stimulus Materials

Table 3-24 summarizes the stimulus materials used in the development of the Speaking Tests. Stimulus materials for both the Speaking (Written) and the Speaking (Oral) Tests were selected

Table 3-24

Speaking Pilot Test Scripts and Sources, Grades 3, 6, 9, 12

Grade	Subtest	Content	Source	Language
3	S3-1	M. M. Green, "Aeroplane"	Mary C. Austin and Queenie B. Mills, (eds.), The Sound of Poetry, 1963. Allyn & Bacon Canada Ltd., 791 St. Clair Avenue West, Toronto M6C 1B8	Poetry Written
	S3-2	"Cats, Cats, Cats"	Unknown	Prose Written
	S3-3	M. M. Green, "Aeroplane"	Mary C. Austin and Queenie B. Mills, (eds.), The Sound of Poetry, 1963. Allyn & Bacon Canada Ltd., 791 St. Clair Avenue West, Toronto M6C 1B8	Poetry Spoken
	S3-4	"Cats, Cats, Cats"	Unknown	Prose Spoken
	S3-5	Picture Sequence	Original	Spontaneous
	S3-6	Picture Stimulus	Original	Articulation
6	S6-1	Lew Sarett, "The Wolf Cry"	Lew R. Sarett, The Collected Poems of Lew Sarett, Holt, Rinehart & Winston, 1941, p. 383. Holt, Rinehart & Winston of Canada Ltd., 55 Horner Avenue, Toronto M8Z 4X6	Poetry Written
	S6-2	William Ettridge, "The Olympic Games"	Robert Neelsen (ed.), The Canadian Children's Annual, p. 6. Potlatch Public, 35 Dalewood Crescent, Hamilton, Ontario L8S 4B5	Prose Written
	S6-3	Lew Sarett, "The Wolf Cry"	Lew R. Sarett, The Collected Poems of Lew Sarett, Holt, Rinehart & Winston, 1941, p. 383. Holt, Rinehart & Winston of Canada Ltd., 55 Horner Avenue, Toronto M8Z 4X6	Poetry Spoken
	S6-4	William Ettridge, "The Olympic Games"	Robert Neelsen (ed.), The Canadian Children's Annual, p. 6. Potlatch Public, 35 Dalewood Crescent, Hamilton, Ontario L8S 4B5	Prose Spoken
	S6-5	Picture Sequence	Original	Spontaneous

Table 3-24 (Continued)

Grade	Subtest	Content	Source	Language
9	S9-1	Mary T. Allan, "The Wolverine"	Thelma Lower and Frederick Coswell (Comp.), The Enchanted Land, W. J. Gage Co. Ltd., Scarborough, Ontario 1967.	Poetry Written
	S9-2	William Ettridge, "The Olympic Games"	Robert Neelsen (ed.), The Canadian Children's Annual, p. 6. Potlatch Public, 35 Dalewood Crescent, Hamilton, Ontario L8S 4B5	Prose Written
	S9-3	"Speleology" Scrambled Details	Original	Spontaneous Written
	S9-4	Mary T. Allan, "The Wolverine"	Thelma Lower and Frederick Coswell (Comp.), The Enchanted Land, W. J. Gage Co. Ltd., Scarborough, Ontario 1967.	Poetry Spoken
	S9-5	"Hiking" Scrambled Details	Original	Spontaneous Spoken
12	S12-1	A. J. M. Smith, "The Lonely Land"	Bliss Corman, Lorne Pierce and V. B. Rhodenezes (eds.), Canadian Poetry in English (Canadian Literature Series), 1954. Ryerson Press--McGraw-Hill Ryerson Ltd., 330 Progress Avenue, Scarborough, Ontario M1P 2Z5.	Poetry Written
	S12-2	"Speleology" Scrambled Details	Original	Spontaneous Written
	S12-4	"Hiking" Scrambled Details	Original	Spontaneous Spoken
	S12-3	A. J. M. Smith, "The Lonely Land"	Bliss Corman, Lorne Pierce and V. B. Rhodenezes (eds.), Canadian Poetry in English (Canadian Literature Series), 1954. Ryerson Press--McGraw-Hill Ryerson Ltd., 330 Progress Avenue, Scarborough, Ontario M1P 2Z5.	Poetry Spoken

after a careful examination of literature suited to each grade level. Since the same selections would be used as stimulus materials for both Written and Oral Tests, the following criteria for selection were set up. Materials selected should:

1. fall within the students' "easy-reading" range
2. possess a sufficiently high interest level to stimulate the students;
3. be written in a style which would lend itself to oral reading; and
4. contain sufficient substance to allow for the type of analysis required to test a variety of objectives.

In order to insure equality in the test situation an attempt was made to choose selections of prose and poetry with which students would not be familiar.

To test Linguistic Factors/Sounds at the Grade 3 level a Picture Articulation Test consisting of 30 frames was devised by the Research Team. Criteria for such test construction was found in Johnson, Darley and Spriestersbach (1963).

Stimulus materials for the Speaking (Written) Test at the Grade 3 and 6 levels consisted of a poem and a prose selection. Items framed to test student knowledges were based on these selections. A poem and a piece of prose were selected as stimulus materials for the Speaking (Written) Test at the Grade 9 level. In addition, a scrambled outline in the form of a set of details abstracted from a prepared speech, was used as the stimulus to elicit student knowledges of organizational factors. At the Grade 12 level a selection of poetry and a scrambled outline became the stimulus materials.

The topic of "Speleology" was chosen for the scrambled outline for both Grades 9 and 12. This topic was seen to be sufficiently unfamiliar to insure equality of background knowledge. The term "speleology" was defined and a set of randomly ordered details prepared. Items were framed to assess the students' ability to perceive relationships among sub-ideas which support a central idea and to test a variety of other organizational knowledges.

Speaking (Oral) stimulus materials at the Grade 3 and 6 levels consisted of the prose and poetry selections previously analyzed in the Written Test. These were later used in the testing of oral reading. For the spontaneous speaking task at these levels the Research Team constructed picture sequence sets, a different set for each grade. In preparing these picture sequences an attempt was made to include high-interest material which would stimulate a variety of responses.

At the Grade 9 and 12 levels the stimulus material for oral reading consisted of the poem previously analyzed in the written test. Stimulus material for spontaneous prepared speech consisted of a set of randomly ordered details in the topic of "Hiking."

Table 3-25 summarizes the Pilot Test items on the stimulus subtests, stimulus format, language type and indicates the number of items in each test and subtest. Tables 3-26 to 3-29 provide the test blueprints for speaking. An Objectives/Items matrix for all grades appears in Table 3-30.

Table 3-25

Speaking Pilot Test Content by Stimulus Subtest, Stimulus Format, Language Type, Item Numbers, and Number and Percent of Items

Stimulus Subtest	Stimulus Format	Language Type	Item Numbers	Total Number of Items	Percent of Items
Grade 3					
1	Poetry	Written	1-5	5	7
2	Prose	Written	6-10	5	7
3	Poetry	Oral	11-22	12	17
4	Prose	Oral	23-34	12	17
5	Picture Sequence	Spontaneous Speech	35-70	36 70	52 100
Grade 6					
1	Poetry	Written	1-9	9	12
2	Prose	Written	10-15	6	8
3	Poetry	Oral	16-26	11	14
4	Prose	Oral	27-38	12	16
5	Picture Sequence	Spontaneous Speech	39-76	38 76	50 100
Grade 9					
1	Poetry	Written	1-9	9	10
2	Prose	Written	10-17	8	9
3	Scrambled Details	Spontaneous Written	18-25	8	9
4	Poetry	Oral	26-39	14	14
5	Given Topic	Planned Spontaneous Oral	40-94	55 94	58 100
Grade 12					
1	Poetry	Written	1-11	11	12
2	Scrambled Details	Spontaneous Written	12-19	8	9
3	Poetry	Oral	20-33	14	16
4	Given Topic	Planned Spontaneous Oral	34-89	56 89	63 100

Table 3-26

Speaking Pilot Test Blueprint, Grade 3

Domain Number	Domain	Obj. #	# Items Per Obj.	# Obj. In Domain	% Obj. In Domain	# Items In Domain	% Items In Domain
1	Voice Production Factors	1A1 1C1 1C2 1D1 1D2 1E3 1E4 1E5	4 2 3 1 3 6 6 1	8	32	26	38
2	Nonverbal Factors	2B1 2B2	1 3	2	9	4	6
3	Linguistic Factors - Sounds	3A1 3A2	3 3	2	8	6	9
4	Linguistic Factors - Words	3B1 3B2 3B6 3B7	5 5 2 1	4	16	13	19
5	Linguistic Factors - Sentences	3C1 3C2 3C9	1 1 2	3	12	4	6
6	Organizational Factors - Ideational Fluency	4A1	1	1	4	1	1
7	Organizational Factors - Plan Analysis	---					
8	Organizational Factors - Organization of details	4C1 4C2 4C4 4C5 4C6	2 1 1 8	5	20	14	21
	Totals	25	68	25	100	68*	100

* This test contains two items which measured general ability in oral reading for which no objective was stated and which therefore do not appear on this list.

Table 3-27

Speaking Pilot Test Blueprint, Grade 6

Domain Number	Domain	Obj. #	# Items Per Obj.	# Obj. In Domain	% Obj. In Domain	# Items In Domain	% Items In Domain
1	Voice Production Factors	1A1	4	11	37	31	42
		1A3	2				
		1A4	1				
		1C1	2				
		1C2	3				
		1D1	2				
		1D2	3				
		1E1	2				
		1E3	6				
		1E4	5				
		1E5	1				
2	Nonverbal Factors	2B1	1	2	7	4	5
		2B2	3				
3	Linguistic Factors - Sounds	3A1	3	2	7	6	8
		3A2	3				
4	Linguistic Factors - Words	3B1	1	4	13	10	14
		3B2	5				
		3B6	2				
		3B7	2				
5	Linguistic Factors - Sentences	3C2	2	5	17	10	14
		3C4	2				
		3C7	4				
		3C8	1				
		3C9	1				
6	Organizational Factors - Ideational Fluency	4A1	1	1	3	1	1
7	Organizational Factors - Plan Analysis	4B5	2	1	3	2	2
8	Organizational Factors - Organization of details	4C2	1	4	13	10	14
		4C4	1				
		4C5	6				
		4C6	2				
	Totals	30	74	30	100	74*	100

* This test contains two items which measured general ability in oral reading for which no objective was stated and which therefore do not appear on this list.

Table 3-28

Speaking Pilot Test Blueprint, Grade 9

Domain Number	Domain	Obj. #	# Items Per Obj.	# Obj. In Domain	% Obj. In Domain	# Items In Domain	% Items In Domain
1	Voice Production Factors	1A1	3	13	30	26	28
		1A3	1				
		1B1	1				
		1C1	2				
		1C2	2				
		1D1	2				
		1D2	2				
		1E1	2				
		1E2	3				
		1E3	4				
		1E4	1				
		1E5	1				
		1E6	2				
2	Nonverbal Factors	2A1	1	5	12	6	6
		2A2	1				
		2B1	1				
		2B2	2				
		2B3	1				
3	Linguistic Factors - Sounds	3A1	2	1	2	2	2
4	Linguistic Factors - Words	3B2	5	5	12	12	13
		3B3	1				
		3B4	1				
		3B6	1				
		3B7	4				
5	Linguistic Factors - Sentences	3C1	6	8	19	24	26
		3C2	1				
		3C4	4				
		3C5	6				
		3C6	1				
		3C8	3				
		3C9	2				
		3C10	1				
6	Organizational Factors - Ideational Fluency	4A1	1	1	2	1	1
7	Organizational Factors - Plan Analysis	4B3	1	3	7	7	7
		4B5	3				
		4B6	3				
8	Organizational Factors - Organization of details	4C1	2	7	16	16	17
		4C2	2				
		4C3	1				
		4C4	4				
		4C5	3				
		4C6	1				
		4C9	3				
	Totals	43	94	43	100	94	100

Table 3-29

Speaking Pilot Test Blueprint, Grade 12

Domain Number	Domain	Obj. #	# Items Per Obj.	# Obj. In Domain	% Obj. In Domain	# Items In Domain	% Items In Domain
1	Voice Production Factors	1A1	3	12	32	26	29
		1A4	2				
		1C1	2				
		1C2	2				
		1D1	2				
		1D2	2				
		1E1	1				
		1E2	4				
		1E3	4				
		1E4	1				
		1E5	1				
		1E6	2				
2	Nonverbal Factors	2A1	2	5	13	7	8
		2A2	1				
		2B1	1				
		2B2	2				
		2B3	1				
3	Linguistic Factors - Sounds	3A1	1	1	3	1	1
4	Linguistic Factors - Words	3B2	5	3	8	9	10
		3B3	1				
		3B7	3				
5	Linguistic Factors - Sentences	3C1	7	7	18	24	27
		3C2	1				
		3C4	4				
		3C5	6				
		3C8	3				
		3C9	2				
		3C10	1				
6	Organizational Factors - Ideational Fluency	---					
7	Organizational Factors - Plan Analysis	4B3	1	4	10	7	8
		4B4	1				
		4B5	2				
		4B6	3				
8	Organizational Factors - Organization of details	4C1	2	6	16	15	17
		4C2	2				
		4C3	1				
		4C4	4				
		4C5	3				
		4C9	3				
	Totals	38	89	38	100	89	100

Table 3-30

Speaking Pilot Test Objectives/Items Matrices,
Grades 3, 6, 9, 12

Domain #	Domain	Obj. #	Grade 3	DT	Grade 6	DT	Grade 9	DT	Grade 12	DT
1	Voice Production Factors N = 14	1A1	4		4		3		3	
		1A3	0		2		1		0	
		1A4	0		1		0		2	
		1B1	0		0		1		0	
		1C1	2		2		2		2	
		1C2	3		3		2		2	
		1D1	1		2		2		2	
		1D2	3		3		2		2	
		1E1	0		2		2		1	
		1E2	0		0		3		4	
		1E3	6		6		4		4	
		1E4	6		5		1		1	
		1E5	1		1		1		1	
		1E6	0	26	0	31	2	26	2	26
2	Nonverbal Factors N = 5	2A1	0		0		1		2	
		2A2	0		0		1		1	
		2B1	1		1		1		1	
		2B2	3		3		2		2	
		2B3	0	4	0	4	1	6	1	7
3	Linguistic Factors - Sounds N = 2	3A1	3		3		2		1	
		3A2	3	6	3	6	0	2	0	1
4	Linguistic Factors - Words N = 6	3B1	5		1		0		0	
		3B2	5		5		5		5	
		3B3	0		0		1		1	
		3B4	0		0		1		0	
		3B6	2		2		1		0	
		3B7	1	13	2	10	4	12	3	9
5	Linguistic Factors - Sentences N = 9	3C1	1		0		6		7	
		3C2	1		2		1		1	
		3C4	0		2		4		4	
		3C5	0		0		6		6	
		3C6	0		0		1		0	
		3C7	0		4		0		0	
		3C8	0		1		3		3	
		3C9	2		1		2		2	
		3C10	0	4	0	10	1	24	1	24
6	Ideational Fluency N = 1	4A1	1	1	1	1	1	1	0	0
7	Organizational Factors N = 4	4B3	-		0		1		1	
		4B4	-		0		0		1	
		4B5	-		2		3		2	
		4B6	-		0	2	3	7	3	7
8	Organizational Factors - Details N = 7	4C1	2		0		2		2	
		4C2	1		1		2		2	
		4C3	0		0		1		1	
		4C4	1		1		4		4	
		4C5	8		6		3		3	
		4C6	2		2		1		0	
		4C9	0	14	0	10	3	16	3	15

Administration Time

The administration time in minutes for the Pilot Speaking Tests at all grade levels is shown in the charts below.

Administration Time - Speaking (Written)

Grade	Distribution of Materials	Testing Time	Total Time
3	10 minutes	30 minutes	40 minutes
6	10 minutes	30 minutes	40 minutes
9	5 minutes	35 minutes	40 minutes
12	5 minutes	35 minutes	40 minutes

Administration Time - Speaking (Oral)

Grade	Distribution of Materials	Preparation	Presentation (Approximately)	Total Time
3	5 minutes	---	6 students x 10 minutes per student = 60 minutes	65 minutes
6	5 minutes	---	6 students x 10 minutes per student = 60 minutes	65 minutes
9	5 minutes	45 minutes per group	6 students x 10 minutes per student = 60 minutes	110 minutes
12	5 minutes	45 minutes per group	6 students x 10 minutes per student = 60 minutes	110 minutes

Administration Manual (Teacher's Manual)

An Administration Manual (Teacher's Manual) was developed for the Speaking Test for all grade levels. Each manual contained two sections. Section I contained instructions for the administration of the Speaking (Written) Test, and was divided into two subtests at the Grade 3, 6, and 12 levels, and three subtests at the Grade 9 level. The format took into account the possible future use of these tests by making each subtest a unit which could be administered separately. For this reason, instructions for the general test administration were repeated at the beginning of each subtest.

The second section, Speaking (Oral) consisted of instructions for the administration of a test of the oral reading of prose and/or poetry and one for spontaneous speech at all grade levels. At the Grade 3 level instructions for the administration of the Articulation Test were also included.

Both the Speaking (Written) and the Speaking (Oral) sections of the Manual-contained reprints of the material presented in the Student Booklet.

Student Test Booklets

The Student Test Booklets contained the same divisions as did the Teacher's Manual. Explanations and instructions preceded each subtest. The written subtests took the form of multiple choice items. At the Grade 3 and 6 levels, students' answers were to be answered in the Test Booklets. At the Grade 9 and 12 levels, students were instructed to place their answers on separate answer sheets which were provided along with the Test Booklets.

ATTITUDE MEASURE

A number of instruments designed to measure attitudes toward communication were examined. A final choice of two instruments designed to measure communication apprehension was made. These tests were seen to measure that attitudes which the literature indicated were most important to human communication. They also satisfied the time constraints of the total test situation.

The Personal Report of Communication Fear (PRCF) was chosen for administration at the Grade 3 and 6 levels. This test consisted of 12 items measured on a five-point scale.

The Personal Report of Communication Anxiety (PRCA), was chosen for administration at the Grade 9 and 12 levels. This test consisted of 25 items measured on a five-point scale.

Instructions for the administration of these tests were included in the Teacher's Manual. Scoring procedures were not included in this draft.

SPEAKING: DEVELOPMENT OF THE FIELD TESTS

Revisions to both the Speaking (Written) and the Speaking (Oral) Tests at each grade level were based on suggestions from teachers and students, observations of project monitors, and statistical data analyses.

In general, the changes in the Speaking (Written) Tests aimed at: (a) reducing testing time; (b) rewording options for clarity where appropriate; and (c) combining or eliminating items where necessary. All of the domains represented in the Pilot Written Test were retained for the Field Written Test; however, in some cases objectives were combined and/or, where objectives related to a number of items, some were deleted.

In the Grade 3 test a change in test administration was made at the suggestion of classroom teachers. In order to accommodate students whose reading levels varied, the test administrators were to read aloud stimulus materials while the student read them silently. Revisions to the Speaking (Oral) Test were based on the same criteria as those applied to the Speaking (Written) Test.

In general, these changes were aimed at: (a) reducing testing time; and (b) eliminating overlapping criteria.

At the Grade 3 and 6 levels the test was shortened by eliminating the oral reading of the poem. This was done because the data did not reveal a sufficient degree of discrimination between the two tasks to warrant the inclusion of both prose and poetry interpretation.

At the Grade 9 and 12 levels the time for speech preparation was cut from 45 to 15 minutes since it was found that the longer preparation period encouraged the students to write out their speeches in full. This mitigated against the use of spontaneous speech which was the stated purpose of the test.

At all grade levels certain items (such as projection) were cut from the second rating scale as it appeared unnecessary to test them more than once. Other items were eliminated because they proved too difficult for raters to assess in a time-limited speech. Rating scales at all levels were improved by placing all items to be evaluated on a single page for ease in handling and by changing the evaluation categories from "Yes/No" to a five-point scale ranging from (1) excellent to (5) poor.

Table 3-31 summarizes the number and percentage of Field Test items by language type. Stimulus materials and sources for all speaking tests are described in Table 3-32. Table 3-33 shows how items were classified on the basis of subtests, format and language type. Tables 3-34 to 3-37 provide the Speaking Test Blueprints for each grade. Table 3-38 presents the total Objectives/Items Matrices.

Table 3-31

Speaking Field Test Items by Grade and Language Type

Grade	Language Type	Total No. of Items	Percentage of Items on Test
3	Written	10	26
	Spontaneous	18	48
	Oral Reading	10	26
	Total	38	100
6	Written	15	33
	Spontaneous	20	44
	Oral Reading	10	23
	Total	45	100
9	Written	23	43
	Spontaneous	22	40
	Oral Reading	9	17
	Total	54	100
12	Written	19	37
	Spontaneous	9	18
	Total	51	100

Table 3-32

Speaking Field Test Scripts and Sources,
Grades 3, 6, 9, 12

Grade	Subtest	Content	Source	Language
3	S3-1	Mary McBride Green, "Airplane" (Aeroplane).	Mary C. Austin and Queenie B. Mills, (eds.), The Sound of Poetry, 1963. Allyn & Bacon Canada Ltd., 791 St. Clair Avenue West, Toronto M6C1B8	Poetry Written
	S3-2	"Cats, Cats, Cats"	Unknown	Prose Written
	S3-3	Picture Sequence	Original	Spontaneous
	S3-4	"Cats, Cats, Cats"	Unknown	Prose Spoken
	S3-6	Picture Stimulus	Original	Articulation
6	S6-1	Lew R. Sarett, "The Wolf Cry."	Lew R. Sarett, The Collected Poems of Lew Sarett, Holt, Rinehart & Winston, 1941, p. 383. Holt, Rinehart & Winston of Canada Ltd., 55 Horner Avenue, Toronto M8Z 4X6	Poetry Written
	S6-2	William Ettridge, "The Olympic Games."	Robert Neelsen (ed.), The Canadian Children's Annual, p. 6. Potlatch Public, 35 Dalewood Crescent, Hamilton, Ontario L8S 4B5	Prose Written
	S6-3	Picture Sequence	Original	Spontaneous
	S6-4	William Ettridge, "The Olympic Games."	Robert Neelsen (ed.), The Canadian Children's Annual, p. 6. Potlatch Public, 35 Dalewood Crescent, Hamilton, Ontario L8S 4B5	Prose Spoken

Table 3-32 (Continued)

Grade	Subtest	Content	Source	Language
9	S9-1	Mary Thomson Allan, "The Wolverine."	Thelma Lower and Frederick Cogswell (Comp.), The Enchanted Land, W. J. Gage Co. Ltd., Scarborough, Ontario, 1967.	Poetry Written
	S9-2	William Ettridge, "The Olympic Games."	Robert Neelsen (ed.), The Canadian Children's Annual, p. 6. Potlatch Public, 35 Dalewood Crescent, Hamilton, Ontario L8S 4B5	Prose Written
	S9-3	"Speleology" Scrambled Details	Original	Spontaneous Written
	S9-4	"Walking" or "Hiking" Scrambled Details	Original	Spontaneous Spoken
	S9-5	Mary Thomson Allan, "The Wolverine."	Thelma Lower and Frederick Cogswell (Comp.), The Enchanted Land, W. J. Gage Co. Ltd., Scarborough, Ontario, 1967.	Poetry Spoken
12	S12-1	Arthur James Marshall Smith, "The Lonely Land."	Bliss Corman, Lorne Pierce and V. B. Rhodenezes, (eds.), Canadian Poetry in English (Canadian Literature Series), 1954. Ryerson Press--McGraw-Hill Ryerson Ltd., 330 Progress Avenue, Scarborough, Ontario M1P 2Z5	Poetry Written
	S12-2	"Speleology" Scrambled Details	Original	Spontaneous Written
	S12-3	"Walking" or "Hiking" Scrambled Details	Original	Spontaneous Spoken
	S12-4	Arthur James Marshall Smith, "The Lonely Land."	Bliss Corman, Lorne Pierce and V. B. Rhodenezes, (eds.), Canadian Poetry in English (Canadian Literature Series), 1954. Ryerson Press--McGraw-Hill Ryerson Ltd., 330 Progress Avenue, Scarborough, Ontario M1P 2Z5	Poetry Spoken

Table 3-33

Speaking Field Test Content by Stimulus Subtest,
Stimulus Format, Language Type, Item Number,
and Number and Percent of Items

Grade 3

Stimulus Subtest	Stimulus Format	Language Type	Item Numbers	Total Number of Items	Percentage of Items
1	Poetry	Written	1-5	5	13
2	Prose	Written	6-10	5	13
3	Picture Sequence - Story	Spontaneous Speech	11-28	18	48
4	Prose	Oral	29-37*	10	26
				38	100
**	Individual Pictures	Articulation (Oral)	38-53	(16)	--

* One item not numbered.
** The Articulation Test is not included in these tables. It Consisted of 16 Items based on speech sounds, Domain #3, Linguistic Factors, Sounds.

Grade 6

1	Poetry	Written	1-9	9	20
2	Prose	Written	10-15	6	13
3	Picture Sequence - Story	Spontaneous Speech	16-35	20	44
4	Prose	Oral	36-44*	10	23
				45	100

* One item not numbered.

Grade 9

1	Poetry	Written	1-9	9	17
2	Prose	Written	10-16	7	13
3	Speech Preparation (Organization)	Written	17-23	7	13
4	A Given Topic	Planned Spontaneous	24-45	22	40
5	Poetry	Oral	46-54	9	17
				54	100

Grade 12

1	Poetry	Written	1-11	11	22
2	Speech Preparation (Organization)	Written	12-19	8	16
3	A Given Topic	Planned Spontaneous	20-42	23	45
4	Poetry	Oral	43-51	9	17
				51	100

Table 3-34

Speaking Field Test Blueprint, Grade 3

Domain Number	Domain	Obj. #	# Items Per Obj.	# Obj. In Domain	% Obj. In Domain	# Items In Domain	% Items In Domain
1	Voice Production Factors	1A1	1	8	35	11	29
		1A2	2				
		1C1	1				
		1D1	1				
		1D2	1				
		1E2	2				
		1E3	2				
		1E4	1				
2	Nonverbal Factors	2B1	1	2	9	3	8
		2B2	2				
3	Linguistic Factors - sounds	3A1	3	2	9	4	10.5
		3A2	1				
4	Linguistic Factors - Words	3B2	2	3	13	5	13
		3B5	2				
		3B6	1				
5	Linguistic Factors - Sentences	3C7	3	2	9	4	10.5
		3C8	1				
6	Organizational Factors - Ideational Fluency	4A1	1	1	4	1	3
7	Organizational Factors - Plan Analysis	---					
8	Organizational Factors - Organization of details	4C1	2	5	21	10	26
		4C2	1				
		4C4	1				
		4C5	4				
		4C6	2				
	Totals	23	38	23	100	38	100

Table 3-35

Speaking Field Test Blueprint, Grade 6

Domain Number	Domain	Obj. #	# Items Per Obj.	# Obj. In Domain	% Obj. In Domain	# Items In Domain	% Items In Domain
1	Voice Production Factors	1A1	1	11	39	17	40
		1A3	2				
		1A4	1				
		1C1	2				
		1C2	2				
		1D1	2				
		1D2	1				
		1E1	2				
		1E2	1				
		1E3	2				
		1E4	1				
2	Nonverbal Factors	2B1	1	2	7	3	7
		2B2	2				
3	Linguistic Factors - Sounds	3A1	2	2	7	3	7
		3A2	1				
4	Linguistic Factors - Words	3B2	2	3	11	6	14
		3B5	2				
		3B6	2				
5	Linguistic Factors - Sentences	3C2	1	4	14	5	12
		3C4	1				
		3C7	2				
		3C8	1				
6	Organizational Factors - Ideational Fluency	4A1	1	1	4	1	2
7	Organizational Factors - Plan Analysis	4B5	2	1	4	2	4
8	Organizational Factors - Organization of details	4C2	1	4	14	6	14
		4C4	1				
		4C5	2				
		4C6	2				
	Totals	28	43*	28	100	43*	100

* This test contains two items which measured general ability in oral reading for which no objective was stated and which therefore do not appear on this list.

Table 3-36

Speaking Field Test Blueprint, Grade 9

Domain Number	Domain	Obj. #	# Items Per Obj.	# Obj. In Domain	% Obj. In Domain	# Items In Domain	% Items In Domain
1	Voice Production Factors	1A1	1	12	35	18	33
		1A3	2				
		1B1	1				
		1C1	1				
		1C2	2				
		1D1	2				
		1D2	1				
		1E1	2				
		1E2	2				
		1E3	2				
		1E4	1				
		1E5	1				
2	Nonverbal Factors	2A2	1	3	9	4	7
		2B3	1				
		2B2	2				
3	Linguistic Factors - Sounds	3A1	2	1	3	2	4
4	Linguistic Factors - Words	3B2	1	4	12	5	9
		3B3	1				
		3B5	1				
		3B6	2				
5	Linguistic Factors - Sentences	3C2	1	6	17.5	9	16
		3C4	1				
		3C5	3				
		3C6	1				
		3C7	2				
		3C8	1				
6	Organizational Factors - Ideational Fluency	---					
7	Organizational Factors - Plan Analysis	4B3	1	2	6	4	8
		4B6	3				
8	Organizational Factors - Organization of details	4C1	1	6	17.5	12	23
		4C2	1				
		4C4	4				
		4C5	3				
		4C6	1				
		4C9	2				
	Totals	34	54	34	100	54	100

Table 3-37

Speaking Field Test Blueprint, Grade 12

Domain Number	Domain	Obj. #	# Items Per Obj.	# Obj. In Domain	% Obj. In Domain	# Items In Domain	% of Items Domain
1	Voice Production Factors	1A1	1	11	34	16	31
		1A4	2				
		1C1	2				
		1C2	1				
		1D1	2				
		1D2	1				
		1E1	1				
		1E2	2				
		1E3	2				
		1E4	1				
		1E5	1				
2	Nonverbal Factors	2A2	1	3	10	4	8
		2B2	2				
		2B3	1				
3	Linguistic Factors - Sounds	3A1 3A2	2	2	6	2	4
4	Linguistic Factors - Words	3B2	1	3	10	5	10
		3B3	1				
		3B6	3				
5	Linguistic Factors - Sentences	3C2	1	5	16	8	16
		3C4	1				
		3C5	3				
		3C7	2				
		3C8	1				
6	Organizational Factors - Ideational Fluency	---					
7	Organizational Factors - Plan Analysis	4B3	1	2	6	3	6
		4B5	2				
8	Organizational Factors - Organization of details	4C1	1	6	18	13	25
		4C2	1				
		4C3	1				
		4C4	4				
		4C5	3				
		4C9	3				
	Totals	32	51	32	100	51	100

Table 3-38

Speaking Field Test Objectives/Items Matrices,
Grades 3, 6, 9, 12

Domain #	Domain	Obj. #	Grade 3	DT	Grade 6	DT	Grade 9	DT	Grade 12	DT
1	Voice Production Factors N = 13	1A1	1		1		1		1	
		1A3	0		2		2		0	
		1A4	0		1		0		2	
		1B1	0		0		1		0	
		1C1	1		1		1		2	
		1C2	2		2		2		1	
		1D1	1		2		2		2	
		1D2	1		1		1		1	
		1E1	0		2		2		1	
		1E2	2		2		2		2	
		1E3	2		2		2		2	
		1E4	1		1		1		1	
		1E5	0	11	0	17	1	18	1	16
2	Nonverbal Factors N = 4	2A2	0		0		1		1	
		2B1	1		1		0		0	
		2B2	2		2		2		2	
		2B3	0	3	0	3	1	4	1	4
3	Linguistic Factors - Sounds N = 2	3A1	3		2		2		2	
		3A2	1	4	1	3	0	2	0	2
4	Linguistic Factors - Words N = 4	3B2	2		2		1		1	
		3B3	0		0		1		1	
		3B5	2		2		1		0	
		3B6	1	5	2	6	2	5	3	5
5	Linguistic Factors - Sentences N = 6	3C2	0		1		1		1	
		3C4	0		1		1		1	
		3C5	0		0		3		3	
		3C6	0		0		1		0	
		3C7	3		2		2		2	
		3C8	1	4	1	5	1	9	1	8
6	Ideational Fluence N = 1	4A1	1	1	1	1	1	1	1	1
7	Organizational Factors N = 3	4B3	0		0		1		1	
		4B5	0		2		0		2	
		4B6	0	0	0	2	3	4	0	3
8	Organizational Factors - Details N = 7	4C1	2		0		1		1	
		4C2	1		1		1		1	
		4C3	0		0		0		1	
		4C4	1		1		4		4	
		4C5	4		2		3		3	
		4C6	2		2		1		0	
		4C9	0	10	0	6	2	12	3	13
	Totals	35	38		43		54		51	

Grade 3 = 23 Objectives; Grade 6 = 28 Objectives; Grade 9 = 34 Objectives; Grade 12 = 32 Objectives.

Speaking Field Test Procedures

The Speaking (Written) Test was administered by regular classroom teachers at all grade levels. Each teacher was provided with: (a) a Teacher's Manual; (b) Student Test Booklets, and (c) Answer Sheets (Grades 6, 9, 12).

The same procedures were followed in the Field Test as in the Pilot Test with the following exceptions. At the Grade 3 level the teacher read aloud the stimulus material while the students read it silently. At the Grade 9 and 12 levels preparation time for the extemporaneous speech was shortened from 45 to 15 minutes in order that a realistic sampling of spontaneous speech could be made.

Administration Time

In the Speaking (Written) Test administration time remained the same in the Field Test as in the Pilot Test.

In the Speaking (Oral) Test administration, the time was shortened at the Grade 3 and 6 levels by the elimination of the oral reading of the poem.

At the Grade 9 and 12 levels the shortening of preparation time for the extemporaneous speech from 45 to 15 minutes reduced the testing time.

The administration time in minutes for both sections of the Field Speaking Test at all grade levels is shown in the charts below.

Administration Time - Speaking (Written) Test

Grade	Distribution	Testing Time	Total Time
3	10 minutes	30 minutes	40 minutes
6	10 minutes	30 minutes	40 minutes
9	5 minutes	35 minutes	40 minutes
12	5 minutes	36 minutes	40 minutes

Administration Time - Speaking (Oral) Test

Grade	Distribution	Preparation	Testing Time	Total Time
3	5 minutes	--	6 x 5 minutes	35 minutes
6	5 minutes	--	6 x 5 minutes	35 minutes
9	5 minutes	15 minutes	6 x 7 minutes	62 minutes
12	5 minutes	15 minutes	6 x 7 minutes	62 minutes

LISTENING AND SPEAKING: FINAL TESTS

Following analysis of Field Test data it was necessary to determine what changes should be made in the production of the Final Test. The results of the data analysis indicated that only minor changes, largely of an editorial nature, needed to be made. Additionally, there were considerations about how the test should be administered. Suggestions about changes in administration did not affect the actual tests themselves and are considered as possible recommendations in Chapter 5.

Final Listening Test Objectives are found in Appendices C-1 through C-4. Final Speaking Test Objectives are found in Appendices C-5 through C-8. Final Listening Test Objectives/Items Matrices are found in Appendices D-1 through D-4. Final Speaking Test Objectives/Items Matrices are found in Appendices D-5 through D-8.

Chapter 4

DATA ANALYSIS AND DISCUSSION

THE PILOT TESTS

In accordance with the plan outlined by the Research Team in their Proposal and approved by the Steering Committee, the Pilot Listening and Speaking Tests were administered to intact Grade 3, 6, 9, and 12 classes in urban and rural schools in the Calgary area during the week of May 8, 1978. The following purposes of the pilot testing were outlined in the Proposal (page 15).

1. To determine if the subjects can understand the directions;
2. to determine if the administrative procedures are adequate;
3. to determine if the options in multiple choice items are working properly;
4. to determine the amount of time required for each item or group of items;
5. to determine if items discriminated or were too difficult or too easy (floor effects and ceiling effects);
6. to determine the practicality and reliability of the scoring system;
7. to provide baseline data for the preparation of the field test instruments.

The Research Team noted in the Proposal that the pilot testing activity was particularly important in the area of speaking, where it was assumed that much less has been done generally in defining objectives and in developing instruments to evaluate attitudes, abilities and skills based on these objectives. (This assumption was supported by the review of the literature.)

Sample

The Pilot Test sample was the same for both the Listening and Speaking Test administrations. The sample consisted of the schools and classes listed below. All contacts with these schools were made by members of the Research Team.

Grade 3: (N=117)

Calgary: Calgary Board of Education, Calgary, Alberta.

Marion Carson Elementary School
5225 Varsity Drive N. W., Calgary
Principal: George Sackman
Two intact classes (N=57)

Okotoks: Foothills School Division #38, Alberta.

Percy Pegler School
Box 700, Okotoks
Principal: David Ballard
Two intact classes (N=60)

Grade 6: (N=124)

Calgary: Calgary Board of Education, Calgary, Alberta.

Marion Carson Elementary School
Two intact classes (N=64)

Okotoks: Foothills School Division #38, Alberta.

Percy Pegler School
Two intact classes (N=60)

Grade 9: (N=101)

Calgary: Calgary Board of Education, Calgary, Alberta.

Milton Williams Junior High School
92 Malibou Road S. W., Calgary
Principal: Keith Hansen
Two intact classes (N=54)

Airdrie: Rockyview School Division #41, Alberta.

Airdrie Junior High School
Box 249, Airdrie
Principal: Alex Semenoff
Two intact classes (N=47)

Grade 12: (N=42)

Calgary: Calgary Board of Education, Calgary, Alberta.

Crescent Heights Senior High School
1019 First Street N. W., Calgary
Principal: Les Roberts
English 30/33 students (N=42)

Springbank: Rockyview School Division #41, Alberta.

Springbank Community High School
R. R. #2, Calgary
Principal: John Znider
English 30/33 students (N=40)

Due to certain administrative difficulties encountered during the Springbank testing, the data obtained from the 40 students tested at Springbank High School could not be used. Thus the total Pilot Test sample at Grade 12 consisted of the students at Crescent Heights High School, Calgary.

PILOT TEST ADMINISTRATION

After selecting the pilot schools, members of the Research Team met with the principals, coordinators, and English department heads in the cooperating schools. Testing schedules were developed during these meetings in accordance with the timetables of each school. The test administration for both Listening and Speaking was done by certified substitute teachers trained by the Research Team. The schools included both closed and open areas, and test administration was conducted under both types of situations.

In general, the Listening Tests were administered first in each school. The substitute teachers who administered the tests were accompanied by at least one member of the Research Team. Materials were brought to the schools by the Research Team and Research Assistants and were collected afterwards by them. All scoring was done at the University of Calgary. Both teacher and student comments concerning the test and their administration were elicited from and freely given by the classroom teachers.

The answer sheets were coded by the Research Assistants for computer scoring. Coded information included the type of test (Listening, Speaking, etc.), the grade level, the school, the class, a student number, and the sex of the students.

Since the administration of the Listening and Speaking Tests were different, the procedure for each is discussed separately below.

Pilot Listening Tests: Procedures

Each student was provided with a Student Test Booklet and an HB pencil. Grade 3 students marked their answers in the Test Booklets. These answers were later transferred to answer sheets. Grade 6, 9, and 12 students marked their answers directly on the answer sheets. Only the options appeared in the Test Booklets. The stimulus materials and the questions were not provided in the Booklets.

All of the Pilot Listening Subtests were tape recorded. With the exception of some introductory comments concerning the nature of the test, and provided by the classroom teachers, all of the information concerning the testing procedures was provided on reel-to-reel audio tapes. The stimulus materials, questions, and options were also provided on the tape. The administration was monitored by one member of the Research Team, two Research Assistants, and the classroom teachers.

In some schools the two classes receiving the test were combined. In others the test was administered in individual classrooms. Sound quality was generally good in all cases.

Pilot Speaking Tests: Procedures

General Procedures. The first part of the speaking test contained a set of Written Subtests. For these, each student was provided with a Student Test Booklet and an HB pencil. Grade 3 students marked their answers in the Test Booklets. These answers were later transferred to the answer sheets. Grade 6, 9, and 12 students marked their answers directly on answer sheets. All stimulus materials, questions, and options were provided in the Test Booklets.

At each grade level the Written Speaking Test included prose and poetry selections which were printed in the Test Booklets. Questions were directed toward eliciting student knowledge about the oral interpretation of literature. At the Grade 9 and 12 levels, an additional test was included which was concerned with eliciting student knowledge about organizing an extemporaneous speech in terms of audience, situation, and purpose.

An Attitude Test was also administered at each grade level at the same time as the Written Speaking Test.

The second section contained an Oral Test, for which administration procedures are described in detail in the next section.

Procedures: Oral Speaking Tests

Grades 3 and 6

1. Students were tested in small groups of approximately 4-6 students. Although it was requested that testing be done in closed areas in the school, in some cases testing was done in open areas due to lack of facilities in the school.
2. Testers presented the stimulus material (a poem, a prose selection, and a set of pictures) to each student on an individual basis. The poem was presented first. Students were asked to read it silently and then to read it orally in the group. The reading was tape recorded. Testers evaluated the live reading using General Checklist A for Oral Reading.
3. The same procedure was followed for the testing of the prose selection. General Checklist B for Oral Reading was used.
4. Next, the students were shown a set of four pictures about which they were asked to tell a story which would have a beginning, a middle, and an ending. Students were allowed two minutes to present the story orally. The tester evaluated the live presentation using Checklist C for Spontaneous Speech. All members of the group were provided with a four-frame answer sheet. They were asked to write (in the appropriate frame) comments about the story they had just heard while the student to speak next was examining the story board and preparing his talk.

5. Next the articulation test was administered individually. The students responded orally to a set of stimulus items (pictures). This test was not tape recorded.

6. The tape recorded stories were later evaluated by two testers using Specific Checklist D for Spontaneous Speech. A random sampling of all tape-recorded selections was later evaluated by two teachers for the purpose of establishing rater reliability. On those occasions when tapes were inaudible suitable substitutions were made.

Grades 9 and 12

1. Students were given 45 minutes to prepare a three-minute extemporaneous speech for which supporting details and materials were provided in the test booklet. They were provided with two worksheets. They were instructed to complete Worksheet #1 indicating: (a) the purpose of their speech; (b) the audience they chose to speak to; (c) the main idea of their speech; and (d) the language level they deemed suitable for the audience and purpose chosen. This sheet was presented to the tester at the time of delivery. Worksheet #2 was used to prepare an outline of the speech and to be used as notes during delivery. Students completed this task in large groups.

2. Groups of 4-6 students were then assembled with a teacher-tester. They were asked to present the three-minute extemporaneous speech they had prepared. No time lapse was allowed between the preparation and the presentation of the speech. The speech was tape recorded and analyzed live by the tester using General Checklist B for Spontaneous Speech.

3. Next, students were instructed to study the poem previously analyzed in the Written Speaking Test and prepare to read it aloud. The poem was then presented to the group and analyzed live by the tester using Checklist A for Oral Reading.

4. Tape recordings of the speech were later evaluated by two teachers using Checklist D for Spontaneous Speech. As in Grades 3 and 6 a random sampling of the recorded selections was evaluated by two teachers for the purpose of establishing rater reliability. On those occasions when tapes were inaudible suitable substitutions were made.

PILOT TEST RESULTS

As indicated earlier there were a number of purposes associated with the administration of the Pilot Tests. Statistical analyses were performed to determine: a) if the options in the multiple choice items were working properly; b) if the items were too difficult or too easy; c) if the items discriminated among students; and d) if the ratings and tests themselves were reliable. In order to see to what extent the Pilot Tests met these criteria, items analyses, choice distributions, and inter-rater reliability coefficients were calculated for each of the tests. The results of the data analyses are discussed below.

Listening. Table 4-1 presents item data for the Grade 3 Listening Pilot Test. We began with a 55-item pool from which a sub-set of items was to be selected which would meet the requirements of length, difficulty, discrimination, and reliability but would also measure important objectives (content validity). It was decided that items would be acceptable if the Difficulty Index was between .30 and .80;

Table 4-1

Item Analysis Data - Grade 3 Listening (Pilot Test) N = 117

Item No.	Diff.	R(Scale)	Item No.	Diff.	R(Scale)
1	.99	.00	31	.48	.33
2	.98	.62	32	.97	.31
3	.97	.56	33	.64	.32
4	1.00	.90	34	.79	.02
5	.95	.75	35	.98	.20
6	.95	.56	36	.51	.29
7	.08	.40	37	.89	.39
8	.96	.25	38	.74	.49
9	.54	.65	39	.96	.31
10	.68	.64	40	.63	.35
11	.34	.47	41	.77	.23
12	.51	.43	42	.90	.28
13	.76	.42	43	.88	.43
14	.28	.55	44	.81	.27
15	.39	.60	45	.76	.42
16	.04	.18	46	.49	.29
17	.91	.21	47	.41	.19
18	.41	.41	48	.71	.43
19	.80	.39	49	.98	.36
20	.96	.23	50	.94	.57
21	.68	.64	51	.96	.60
22	.90	.37	52	.79	.57
23	.42	.48	53	.74	.49
24	.72	.55	54	.40	.59
25	.91	.38	55	.73	.60
26	.65	.35			
27	.88	.67			
28	.93	.45			
29	.92	.49			
30	.70	.38			

if the Discrimination Index, based on Alpha Coefficients for subscales, was above +.20; and if an option or distractor was chosen by not less than five per cent of the group.

Table 4-1 indicates that out of 55 items 26 appear to be too easy. Thus the average difficulty level was only .70. Six of these items were on the first subscale which had to do with sounds. Items such as these suggested that Grade 3 students could perform certain functions at a very high level and that testing of certain objectives would not be necessary. We also note that five items did not reach criterion in terms of discrimination. A few distractors also failed to attract students, and items in which this occurred were also noted. Using these data the Research Team proceeded to select and revise items for the Field Test. The disposition of items has been discussed in detail in Chapter 3.

Table 4-2 presents the Listening Pilot Test Item Analysis Data for Grade 6. Here we note that 11 items were somewhat easy. The average difficulty level was .68. Only three items failed to discriminate effectively. Again the pool of items appeared adequate as a source for the preparation of the Field Test.

Table 4-3 presents similar results for Grade 9. Here we note that the average difficulty level is .57. Only seven items appeared to be too easy and four items failed to discriminate.

In Grade 12 (see Table 4-4) the average difficulty was .55. There were 17 items that appeared to be somewhat easy, but there were also a number of more difficult items. In terms of discrimination 16 items were inadequate. A number of these failed to discriminate because they were too easy.

Table 4-5 provides a summary of the means, standard deviations, and reliabilities for the Pilot Tests in Listening. The means suggest that the tests were reasonably challenging. It was anticipated that the deletion of very simple items and the revision of

Table 4-2

Item Analysis Data - Grade 6 Listening (Pilot Test) N = 114

Item No.	Diff.	R(Scale)	Item No.	Diff.	R(Scale)
1	.55	.51	26	.58	.27
2	.37	.36	27	.61	.21
3	.25	.10	28	.80	.17
4	.39	.38	29	.62	.51
5	.59	.36	30	.65	.32
6	.23	.39	31	.75	.39
7	.11	.43	32	.50	.53
8	.79	.40	33	.59	.65
9	.77	.42	34	.66	.59
10	.64	.31	35	.51	.54
11	.68	.35	36	.76	.38
12	.14	.19	37	.55	.46
13	.73	.53	38	.64	.50
14	.68	.48	39	.40	.50
15	.93	.19	40	.64	.40
16	.75	.45	41	.70	.41
17	.79	.42	42	.93	.43
18	.83	.39	43	.97	.30
19	.79	.53	44	.92	.55
20	.96	.30	45	.68	.59
21	.84	.45	46	.72	.52
22	.82	.64	47	.89	.36
23	.50	[illegible]	48	.36	.52
24	.65	.66	49	.56	.63
25	.34	.36	50	.54	.61

Table 4-3

Item Analysis Data - Grade 9 Listening (Pilot Test) N = 109

Item No.	Diff.	R(Scale)	Item No.	Diff.	R(Scale)
1	.79	.41	31	.77	.56
2	.60	.58	32	.83	.47
3	.93	.52	33	.39	.59
4	.59	.54	34	.74	.51
5	.63	.69	35	.85	.44
6	.55	.16	36	.71	.43
7	.65	.52	37	.06	.08
8	.58	.48	38	.59	.50
9	.71	.33	39	.39	.73
10	.74	.35	40	.23	.45
11	.49	.44	41	.27	.65
12	.36	.41	42	.11	.50
13	.50	.31	43	.19	.60
14	.40	.24	44	.47	.87
15	.41	.30	45	.40	.85
16	.71	.41	46	.72	.30
17	.83	.30	47	.64	.48
18	.13	.37	48	.31	.16
19	.90	.36	49	.49	.48
20	.20	.12	50	.73	.38
21	.46	.22	51	.82	.53
22	.39	.25	52	.19	.29
23	.79	.37	53	.45	.34
24	.27	.32	54	.66	.44
25	.54	.22	55	.86	.31
26	.78	.37	56	.61	.44
27	.68	.31	57	.69	.52
28	.74	.23	58	.77	.36
29	.69	.54	59	.74	.46
30	.71	.68	60	.63	.26

Table 4-4

Item Analysis Data - Grade 12 Listening (Pilot Test) N = 40

Item No.	Diff.	R(Scale)	Item No.	Diff.	R(Scale)
1	.15	.02	31	.70	.64
2	.79	.73	32	.75	.68
3	.45	.65	33	.85	.47
4	.05	.24	34	.80	.68
5	1.00	.00	35	.85	.75
6	.50	.55	36	.95	.41
7	.35	.05	37	.65	.50
8	.60	.68	38	.00	.00
9	.35	.16	39	.79	.57
10	.80	.10	40	.90	.13
11	.55	.63	41	.25	.33
12	.35	.74	42	.60	.71
13	.30	.61	43	[illegible]	.57
14	.40	.56	44	.30	.48
15	.45	.39	45	1.00	.00
16	1.00	.00	46	.95	-.13
17	.65	-.04	47	.85	.49
18	.90	.11	48	.95	.33
19	.30	.11	49	.65	.71
20	.85	.27	50	.80	.30
21	.15	.33	51	.60	.40
22	.40	.35	52	1.00	.00
23	.45	.22	53	1.00	.00
24	.70	.49	54	.85	.61
25	.45	.34	55	.15	.24
26	.50	.34	56	.65	.43
27	.50	.34	57	.30	.46
28	.75	.44	58	.40	.34
29	.45	.16	59	.50	.47
30	.65	.70	60	.95	.29
			61	.65	.51

Table 4-5

Means, Standard Deviations and Reliability Coefficients for Listening Pilot Tests

Grade	No. of Items	Mean	s.d.	Alpha
3	55	39.71	5.31	.74
6	50	31.67	4.88	.61
9	60	33.89	5.28	.58
12	61	33.85	6.11	.76

others would produce a more difficult test. The standard deviations indicated that we were, in fact, spreading students out on a scale, thus making discrimination possible. The reliabilities ranged from .58 to .76. Considering the heterogeneity of test items, this level of reliability appeared acceptable. It was, of course, assumed that these figures would rise if weak items were removed.

In general, then, the statistical data derived from analysis of pilot test results indicated that a viable base existed from which to select and adapt items for the Field Tests in Listening.

Speaking (Written). Tables 4-6 to 4-9 present the Item data for the Written Speaking Tests. The same criteria that were used in the Listening Tests were applied here. It can be seen that most of the items met the criteria of difficulty and discrimination as specified. Where weaknesses appeared, revisions were undertaken. It should be reiterated, however, that a primary purpose of the Written Test was as preparation for some form of speech production. Thus items might still be useful even if they were easy and had to be retained for the Field Test.

Table 4-6

Item Analysis Data - Grade 3 Speaking (Written) (Pilot Test) N = 113

Item No.	Diff.	R(Scale)	Item No.	Diff.	R(Scale)
1	.50	.61	6	.73	.22
2	.83	.68	7	.85	.70
3	.70	.65	8	.68	.70
4	.78	.70	9	.64	.66
5	.43	.58	10	.50	.66

Table 4-7

Item Analysis Data - Grade 6 Speaking (Written) (Pilot Test) N = 112

Item No.	Diff.	R(Scale)	Item No.	Diff.	R(Scale)
1	.56	.25	11	.90	.46
2	.68	.49	12	.65	.45
3	.71	.39	13	.61	.30
4	.49	.36	14	.71	.35
5	.61	.51	15	.79	.45
6	.50	.20			
7	.45	.34			
8	.20	.23			
9	.20	.27			
10	.86	.58			

Table 4-8

Item Analysis Data - Grade 9 Speaking (Written) (Pilot Test) N = 103

Item No.	Diff.	R(Scale)	Item No.	Diff.	R(Scale)
1	.76	.52	16	.94	.59
2	.89	.57	17	.52	.43
3	.79	.52	18	.72	.44
4	.62	.48	19	.65	.49
5	.83	.42	20	.56	.40
6	.49	.50	21	.49	.42
7	.94	.19	22	.50	.29
8	.76	.57	23	.43	.45
9	.75	.34	24	.68	.51
10	.92	.68	25	.53	.43
11	.49	.32			
12	.87	.66			
13	.91	.64			
14	.90	.68			
15	.74	.42			

Table 4-9

Item Analysis Data - Grade 12 Speaking (Written) (Pilot Test) N = 42

Item No.	Diff.	R(Scale)	Item No.	Diff.	R(Scale)
1	.86	.41	11	.38	.05
2	.74	.32	12	.61	.25
3	.88	.47	13	.50	.51
4	.76	.48	14	.60	.32
5	.38	.45	15	.29	.30
6	.74	.48	16	.52	.55
7	.50	.28	17	.71	.59
8	.86	.58	18	.66	.71
9	.79	.34	19	.54	.56
10	.69	.19			

Table 4-10 presents the test analysis data for the Speaking (Written) Tests. Reliability coefficients range from .50 to .73, suggesting considerable homogeneity among items. Means and standard deviations suggested that the tests were of reasonable difficulty and that a spread of scores was being obtained. With some revision it was felt that the tests were in suitable form to proceed to field testing.

Table 4-10

Means, Standard Deviations and Reliability Coefficients For Speaking Pilot Tests (Written)

Grade	No of Items	Mean	s.d.	Alpha
3	10	6.02	2.50	.69
6	15	8.30	2.85	.59
9	25	17.52	3.90	.73
12	19	12.14	3.48	.50

Speaking (Rating Scales)

The Rating Scales used in the Pilot Test had a 0-1 format; that is, raters marked only the presence or absence of indicators of speaking ability. Analysis of these data indicated that there was not sufficient scope for scoring. Raters did not appear to be willing to make such a specific decision. For this reason it was decided that a five-point scale, based on specific rating instructions, would be used in the Field Test.

Rater Reliability

In order to test the reliability of the ratings made by the judges on the Speaking production items, a sample of the taped oral readings and speeches was selected and rated. This was done by selecting every fifth tape in each grade. Two raters--substitute teachers who had administered the tests but who had not rated the students whose tapes were now selected--rated each tape. It should be noted that items which required visual contact between rater and student were not included in this analysis. The results of these ratings were correlated to arrive at reliability figures. Correlation coefficients ranged from .75 to .83. The average correlation was .78 for all four grades.

In order to determine whether there was adequate validity in the ratings we also looked at choice distributions. There appeared to be considerable spread in the ratings indicating that raters did not simply assign the same score to students within any one scale. We did note, however, that there were several items which raters frequently omitted. These items were revised by providing more rating information and were included again in the Field Test because it was felt that they measured important objectives.

We also considered the feasibility of changing the form of the Rating Scales. In the Pilot Test a 1-0 (Yes, No) rating scheme was used. We felt that teachers would be unable, and perhaps unwilling, to use a five-point scale for the number of items that were required. The results of our analysis and comments from the teachers suggested that a five-point scale would be feasible, and so the Rating Scales were adjusted to this format for the Field Test. At the same time we realized that this procedure might possibly reduce the reliability of scoring.

THE FIELD TESTS

In accordance with the plan outlined by the Research Team in their Proposal and approved by the Steering Committee, the Field Listening and Speaking Tests were administered in intact Grade 3, 6, and 9 classes in urban and rural schools in Didsbury and Red Deer during the weeks of June 12-19, 1978 and during the week of September 18 in Grade 12 classes. The testing of the Grade 12 classes in September was the result of the pressures of a province-wide English testing program, school-wide testing programs, and graduation activities in June, which school administrators felt precluded involvement of Grade 12's in the field testing at that time.

The Proposal called for testing of about 200 pupils per test at each grade level. Further, the Proposal suggested that it would be useful to administer both the Listening and Speaking Tests to the same population at the same time, in order to simplify administrative arrangements as well as to provide a better base for statistical analyses.

Sample

The Field Test sample was the same for both the Listening and Speaking Test administration. This sample is described below.

It should be noted that the original plan was to use the Calgary Board of Education schools for the field testing as well as schools in Didsbury, Alberta. However, when it was discovered that the Calgary School Board did not feel it could participate in the field testing so late in the school year, the cooperation of the Red Deer School Division #104 was requested and granted.

Grade 3: (N = 210)

Red Deer: Red Deer School District #104,
Red Deer, Alberta.
G. W. Smith Elementary School
17 Springfield Avenue
Principal: Alice Hogan
Two intact classes (N = 43)

Oriole Park Elementary School
5 Oldsbury Street
Principal: Harry Kuharchuk
Two intact classes (N = 48)

West Park Elementary School
3814 Fifty-fifth Avenue
Principal: Bert Lougheed
Two intact classes plus the Grade 3 group in a split 3/4 class (N = 60)

Didsbury: County of Mountain View School Division #17,
Didsbury, Alberta.
Ross Ford Elementary School
Didsbury, Alberta.
Principal: Brian Large
Two intact classes (N = 59)

Grade 6: (N = 235)

Red Deer: G. W. Smith Elementary School
Two intact classes (N = 49)

Oriole Park Elementary School
Two intact classes (N = 54)

West Park Elementary School
Two intact classes (N - 62)

Didsbury: Ross Ford Elementary School
Three intact classes (N = 70)

Grade 9: (N = 251)

Red Deer: Eastview Junior High School
3929 Fortiet Avenue
Principal: Larry Pizzey
Seven intact classes (several students were missing from each class due to a previously arranged field trip) (N = 168)

Grade 9: (N = 251)

Didsbury: Didsbury High School
Box 700, Didsbury, Alberta.
Principal: Steve Davidchuk
Four intact classes (N = 83)

Grade 12: (N = 168)

This part of the sample was made up of English 30 students only. This was necessary for administration reasons. It should be noted that the Pilot Test revealed no differences between the English 30 and 33 students, so it was felt that this limitation was appropriate.

Red Deer: Lindsay Thurber Comprehensive High School
4204 - 58th Street,
Red Deer, Alberta
Principal: Dave Blacker
English 30 students (N = 108)

Didsbury: Didsbury High School
Box 700, Didsbury, Alberta
Principal: Warren Phillips
English 30 students (N = 60)

FIELD TEST ADMINISTRATION

In accordance with the procedures described in the Proposal (page 18), contact with the cooperating schools was made first by the Project Monitor in consultation with the Research Team. When the schools and classes were established, correspondence was begun with the school administrators, describing the nature and purpose of the study, the nature of required participation, and the procedures to be followed. When all of the schools had agreed, members of the Research Team met with consultants, administrators, and teachers to develop test administration schedules. In most schools, the Listening and Written Speaking Tests were administered prior to the Oral Speaking Tests. In every case, the Written Speaking Tests were administered before the Oral Speaking Tests.

The Speaking (Oral) Tests were administered by certified teachers other than the students' regular teachers, who were trained by the Research Team. (In a number of cases, these people had also administered the Pilot Speaking (Oral) Tests. The Listening and Speaking (Written) Tests were administered to the students by their regular teachers.

During a series of orientation sessions in each participating school, faculty members involved with the testing program were provided with the necessary testing materials as well as with the following information about the testing procedure:

a. purposes of the study and the testing program;
b. preparation required by both students and participating teachers for the test administration;
c. procedures involved in the test administration;
d. procedures for providing student demographic data required for the study;
3. procedures for packaging and returning the test materials.

Following the testing, the Research Assistants transferred the answers marked by the Grade 3 students in their Test Booklets to answer sheets. They then coded all answer sheets (Grades 3, 6, 9, and 12) in accordance with a coding scheme developed by the Research Team. This coding scheme is provided on page 201.

Field Test Coding Guide

Column	Data File	
1	Test Type:	Listening - 1 Speech, Written - 2 General Spontaneous Speech - 3 Specific Spontaneous Speech - 4 General Oral Reading - 5 Attitude - 6
2	Grade:	2(12); 3(3); 6(6); 9(9); 12(1/2)
3	Schools:	1-8 1 Ross Ford Elementary, Didsbury 2 Didsbury Junior High School, Didsbury 3 Oriole Park Elementary, Red Deer 4 West Park Elementary, Red Deer 5 G. W. Smith Elementary, Red Deer 6 Eastview Junior High School, Red Deer 7 Didsbury Senior High School, Didsbury 8 Lindsay Thurber High School, Red Deer
4	Class:	1-9 (by teacher)
5 & 6	Student Number:	01-60
7	Sex:	Male - 0: Female - 1
8-13	Date of Birth:	Year (8 & 9); Month (10 & 11); Date (12 & 13)

The certified substitute teachers who administered the Speaking Tests in Didsbury were provided by the Research Team, and consisted of Calgary teachers who had been trained by the Research Team and who had participated in the pilot testing.

The certified substitute teachers who administered the Speaking Tests in Red Deer were generally provided by the Red Deer School District, although several additional testers were brought from Calgary to supplement the Red Deer group. Special training sessions were conducted with these teachers by Laura Muir of the Research Team.

A special list of materials was prepared for these testers. The Speaking Test list is provided below.

1. Teacher's Manual

2. Set of Checklists
    - A. General Spontaneous Speaking Checklist (for use when listening to students give their talks) (3, 6, 9, 12)
    - B. Specific Spontaneous Speaking Checklist (for use when listening to the tape recorded versions of the students' talks) (3, 6, 9, 12)
    - C. Oral Reading Checklist (for use when listening to students read the required prose or poetry) (6, 9, 12)
    - D. Articulation Test Form (3)
    - E. Explanation of Checklists (3, 6, 9, 12)

3. Student Test Booklets for Oral Spontaneous Speaking (for student use in preparing talks) (9, 12)

4. Laminated Materials
    - A. Picture Sets (3, 6)
    - B. Pieces of Prose and Poetry (3, 6)
    - C. Articulation Test (3)

5. Set of Blank Tapes (for use in recording students' talks)

The Listening Test materials consisted of:

1. Teacher's Manual
2. Student Booklets
3. Answer Sheets
4. Cassette Tape(s)

Members of the Research Team spent several days in Red Deer in June and in September monitoring the testing. Comments were elicited from both teachers and administrators, who were very helpful and honest. The testing in Red Deer was coordinated by Andy Schaufert, Language Arts Consultant for the city schools, who was enormously cooperative and efficient. During the testing, members of the Research Team and Mr. Schaufert were interviewed by the Red Deer Advocate, and a very positive and lengthy article concerning the testing program resulted from this interview. The members of the Research Team felt strongly that they had excellent relationships with the cooperating schools, and appreciated very much the interest and enthusiasm of the faculty members.

In general, the same procedures used in administering the Pilot Listening and Speaking Tests were followed in administering the Field Tests. However, at the request of the schools, the Listening Tests were provided on cassette tapes. In addition, at the suggestion of the Grade 3 pilot teachers the Grade 3 stimulus materials for the Written Speaking Test were read aloud by the teachers at the same time the students read these silently. This was suggested in order to compensate for any differences in student reading levels.

In the following section we deal with the results of the Field Tests and discuss their implications for further test development and research. Revisions to the present Field Test, based on these findings, will be discussed in the next section.

It should be noted, as pointed out earlier, that the Research Team set out to prepare a criterion- rather than a norm-referenced test. The data as presented, therefore, must be viewed in this light. For example, reliability coefficients which reflect homogeneity of test items would be expected to be much lower in a test such as this one because items were constructed to measure a broad range of objectives rather than a sub-set of objectives related to a narrow band of skills or knowledge. Similarly, correlations among sub-tests can be expected to be low. The important issue in a test such as this one must be whether or not the items themselves are adequate measures of the extent to which students are capable of achieving previously agreed-upon criteria. Notwithstanding this argument for presenting statistical data relevant to criterion-referenced testing, we present below various forms of data analyses which may be useful for further test development and for MACOSA's decision-making needs.

ITEM ANALYSIS DATA

Listening

Tables 4-11 to 4-14 present the item analysis data for the Listening Tests in Grades 3, 6, 9, and 12. It will be noted that the information presented consists of the item number, the stimulus sub-scale of which that item is a part, the objective which the item purports to measure, the difficulty index, the point biserial correlation based on the sub-scale and the point biserial

TABLE 4-11

Item Analysis Data - Grade 3 Listening Field Test

			MF = 211			M = 99			F = 112		
ITEM	SCALE	OBJ.	DIFF.	R(TOTAL)	R(SCALE)	DIFF.	R(TOTAL)	R(SCALE)	DIFF.	R(TOTAL)	R(SCALE)
1	1	5C2C	.70	.12	.40	.71	-.08	.41	.70	.29	.38
2	1	5C2C	.78	.29	.56	.82	.26	.44	.74	.32	.64
3	1	5C2F	.51	.17	.61	.45	.09	.57	.55	.22	.64
4	1	5C1C	.73	.32	.56	.72	.33	.50	.73	.31	.61
5	2	5C1A	.60	.37	.49	.60	.38	.38	.60	.36	.60
6	2	5C1E	.67	.42	.57	.66	.39	.57	.68	.43	.56
7	2	5C1E	.71	.25	.56	.70	.23	.54	.70	.26	.55
8	2	5C2F	.65	.09	.37	.60	.06	.31	.69	.11	.40
9	2	5C1H	.66	.20	.42	.62	.09	.45	.70	.29	.39
10	2	5C3D	.79	.23	.44	.73	.19	.37	.85	.27	.51
11	3	2E1	.73	.26	.44	.76	.26	.53	.71	.27	.36
12	3	5C1A	.60	.31	.46	.57	.38	.45	.63	.26	.47
13	3	5C2H	.91	.05	.24	.90	.08	.16	.92	.02	.32
14	3	4B1	.32	.33	.48	.32	.33	.45	.32	.33	.51
15	3	5C1C	.64	.10	.52	.65	.19	.61	.62	.02	[illegible]
16	4	4B1	.75	.37	.47	.72	.48	.51	.77	.26	.43
17	4	5C1E	.80	.41	.44	.81	.35	.40	.77	.43	.46
18	4	4B7	.71	.30	.38	.72	.36	.42	.68	.25	.33
19	4	4B1	.48	.45	.47	.49	.42	.48	.46	.47	.47
20	4	5C30	.81	.26	.23	.82	.27	.24	.79	.26	.24
21	4	2E1	.70	.25	.34	.66	.25	.32	.74	.24	.35
22	4	4B7	.89	.26	.31	.88	.29	.33	.89	.23	.29
23	4	2D1	.69	.36	.40	.73	.29	.35	.66	.43	.45
24	4	5C30	.79	.26	.34	.81	.14	.23	.77	.36	.43
25	4	5C2F	.84	.30	.36	.81	.35	.35	.87	.26	.36

TABLE 4-11 (Continued)

Item Analysis Data - Grade 3 Listening Field Test

			MF = 211			M = 99			F = 112		
ITEM	SCALE	OBJ.	DIFF.	R(TOTAL)	R(SCALE)	DIFF.	R(TOTAL)	R(SCALE)	DIFF.	R(TOTAL)	R(SCALE)
26	4	5C2H	.78	.47	.50	.72	.36	.47	.83	.56	.54
27	4	5C3D	.38	.33	.37	.39	.31	.35	.36	.35	.37
28	4	5C1H	.73	.19	.25	.70	.19	.27	.75	.19	.23
29	4	5C30	.48	.07	.15	.47	.02	.08	.48	.11	.21
30	5	2B1	.63	.36	.59	.60	.30	.59	.63	.40	.57
31	5	2B1	.51	.30	.61	.46	.27	.58	.54	.32	.63
32	5	2D1	.57	.41	.51	.56	.43	.46	.57	.39	.54
33	5	2D1	.87	.27	.55	.85	.35	.64	.88	.17	.40
34	5	2D1	.23	.06	.35	.26	.07	.33	.20	.07	.38
35	5	2D1	.28	-.01	.24	.26	.11	.39	.29	-.11	.08

TABLE 4-12

Item Analysis Data - Grade 6 Listening Field Test

			MF = 225			M = 108			F = 117		
ITEM	SCALE	OBJ.	DIFF.	R(TOTAL)	R(SCALE)	DIFF.	R(TOTAL)	R(SCALE)	DIFF.	R(TOTAL)	R(SCALE)
1	1	4B3	.61	.22	.55	.62	.26	.60	.60	.16	.51
2	1	5C2F	.37	.23	.50	.29	.18	.41	.44	.24	.55
3	1	5C2C	.47	.24	.49	.44	.29	.38	.49	.18	.57
4	1	5C3D	.66	.11	.31	.69	.09	.40	.61	.21	.28
5	1	5C30	.76	.20	.39	.78	.19	.37	.75	.25	.44
6	2	4B1	.78	.19	.40	.84	.28	.43	.71	.18	.42
7	2	5C2F	.84	.23	.40	.81	.32	.46	.84	.24	.45
8	2	5C2C	.72	.12	.45	.79	.04	.40	.64	.27	.51
9	2	5C3D	.62	.24	.43	.58	.14	.37	.65	.33	.49
10	2	2E1	.81	.44	.49	.80	.39	.46	.81	.54	.58
11	2	4B7	.72	.31	.49	.69	.32	.49	.73	.34	.52
12	3	5C1A	.78	.29	.41	.69	.37	.48	.85	.29	.35
13	3	2D1	.67	.20	.56	.67	.07	.50	.68	.32	.60
14	3	5C2H	.70	.31	.53	.69	.28	.56	.69	.37	.50
15	3	4B3	.44	.12	.44	.41	.20	.38	.46	.08	.52
16	3	5C1E	.55	.23	.45	.56	.33	.56	.53	.11	.35
17	4	5C1P	.43	.20	.56	.41	.23	.63	.47	.12	.48
18	4	5C1P	.64	.31	.68	.54	.33	.68	.72	.31	.65
19	4	2B1	.38	.14	.44	.37	.08	.37	.37	.23	.52
20	4	5C1P	.48	.33	.44	.45	.44	.50	.49	.28	.38
21	5	5C1A	.70	.34	.49	.73	.34	.44	.67	.40	.56
22	5	5C1C	.64	.35	.49	.62	.39	.50	.68	.28	.46
23	5	5C1E	.55	.29	.60	.51	.30	.62	.57	.31	.59
24	5	5C1E	.59	.40	.61	.51	.38	.55	.65	.44	.67
25	5	5C1H	.60	.20	.44	.59	.25	.44	.60	.13	.44

TABLE 4-12 (Continued)

Item Analysis Data - Grade 6 Listening Field Test

			MF = 225			M = 108			F = 117		
TEM	SCALE	OBJ.	DIFF.	R(TOTAL)	R(SCALE)	DIFF.	R(TOTAL)	R(SCALE)	DIFF.	R(TOTAL)	R(SCALE)
26	5	5C30	.86	.35	.40	.87	.25	.44	.84	.51	.41
27	6	4B1	.45	.25	.40	.49	.31	.47	.42	.14	.31
28	6	5C1C	.56	.32	.37	.57	.47	.40	.55	.23	.37
29	6	2D1	.45	.21	.38	.43	.11	.33	.46	.29	.43
30	6	2E1	.42	.21	.31	.36	.27	.34	.45	.21	.33
31	6	2B1	.70	.21	.25	.66	.19	.39	.72	.21	.32
32	6	5C30	.40	.28	.37	.46	.39	.40	.35	.18	.34
33	6	5C2F	.81	.30	.42	.83	.27	.40	.77	.38	.43
34	6	5C30	.70	.35	.32	.68	.29	.28	.72	.35	.35
35	6	5C2H	.57	.33	.41	.59	.26	.37	.57	.31	.41
36	6	2B1	.81	.36	.40	.82	.38	.42	.79	.35	.38
37	7	2B1	.65	.29	.53	.66	.22	.52	.62	.37	.52
38	7	4B7	.58	.34	.41	.63	.28	.36	.54	.38	.59
39	7	4B1	.51	.29	.58	.54	.32	.54	.47	.24	.61
40	7	5C1H	.63	.26	.54	.58	.35	.55	.68	.13	.54

TABLE 4-13

Item Analysis Data - Grade 9 Listening Field Test

			MF = 244			M = 108			F = 136		
ITEM	SCALE	OBJ.	DIFF.	R(TOTAL)	R(SCALE)	DIFF.	R(TOTAL)	R(SCALE)	DIFF.	R(TOTAL)	R(SCALE)
1	1	4B1	.67	.24	.54	.65	.27	.56	.70	.21	.52
2	1	4B7	.76	.47	.64	.73	.49	.64	.78	.45	.64
3	1	4C5	.38	.07	.56	.42	.21	.59	.34	-.05	.53
4	1	4C5	.81	.30	.53	.82	.28	.48	.79	.33	.58
5	2	4B3	.63	.22	.35	.67	.19	.37	.60	.28	.37
6	2	4B7	.49	.21	.48	.53	.17	.55	.46	.25	.41
7	2	4C2	.64	.20	.32	.63	.17	.33	.64	.21	.31
8	2	5C2H	.78	.29	.40	.78	.33	.48	.78	.28	.34
9	2	5C1D	.40	.20	.33	.46	.14	.33	.35	.26	.33
10	2	5C1C	.45	.18	.34	.47	.24	.36	.44	.13	.32
11	2	5C2F	.45	.23	.45	.42	.23	.47	.48	.22	.42
12	2	5C3B	.34	.24	.37	.26	.25	.23	.42	.20	.48
13	2	5C1D	.49	.36	.41	.50	.38	.36	.49	.34	.43
14	3	5C1A	.76	.16	.31	.75	.18	.26	.76	.17	.34
15	3	5C2H	.87	.27	.48	.82	.36	.65	.90	.14	.26
16	3	5C1G	.52	.17	.38	.49	.18	.30	.54	.17	.43
17	3	5C1G	.52	.23	.35	.45	.15	.36	.57	.26	.33
18	3	4B7	.86	.27	.34	.84	.41	.48	.89	.11	.19
19	3	5C30	.35	.21	.39	.34	.14	.30	.37	.28	.50
20	3	5C3E	.62	-.10	.37	.63	.06	.44	.61	-.06	.32
21	3	5C2F	.67	.23	.32	.70	.33	.45	.64	.18	.25
22	3	5C2C	.72	.10	.39	.66	.02	.36	.78	.17	.40
23	4	5C1P	.48	.22	.71	.42	.19	.70	.52	.23	.72
24	4	5C1P	.74	.24	.49	.69	.28	.47	.78	.17	.49
25	4	2B1	.26	.08	.53	.30	.06	.62	.22	.13	.48

TABLE 4-13 (Continued)

Item Analysis Data - Grade 9 Listening Field Test

			MF = 244			M = 108			F = 136		
ITEM	SCALE	OBJ.	DIFF.	R(TOTAL)	R(SCALE)	DIFF.	R(TOTAL)	R(SCALE)	DIFF.	R(TOTAL)	R(SCALE)
26	5	5C3E	.68	.16	.36	.74	.14	.40	.63	.19	.33
27	5	5C1A	.64	.21	.35	.58	.21	.31	.68	.17	.38
28	5	5C1C	.79	.32	.40	.77	.40	.50	.80	.25	.31
29	5	5C1H	.70	.34	.46	.67	.35	.41	.72	.30	.48
30	5	5C1G	.27	.08	.26	.28	.17	.35	.25	.00	.17
31	5	5C1G	.49	.27	.45	.48	.34	.53	.51	.23	.40
32	5	5C1E	.64	.29	.47	.58	.29	.52	.69	.29	.43
33	5	5C1E	.94	.17	.30	.93	.08	.26	.95	.26	.34
34	6	5C2C	.48	.34	.63	.36	.32	.59	.55	.33	.64
35	6	2D1	.84	.23	.48	.80	.26	.53	.87	.19	.41
36	6	5C30	.50	.24	.56	.48	.23	.54	.51	.25	.58
37	6	5C30	.74	.27	.53	.72	.28	.55	.77	.26	.54
38	7	4B3	.76	.18	.32	.73	.21	.38	.78	.12	.24
39	7	5C3B	.41	.00	.48	.35	-.14	.43	.45	.09	.51
40	7	5C30	.65	.16	.41	.62	.22	.43	.68	.09	.39
41	7	2B1	.99	.05	.16	.98	.06	.16	.99	.03	.16
42	7	5C1C	.50	.40	.44	.56	.52	.48	.46	.32	.42
43	7	2D1	.36	.09	.39	.42	.13	.44	.33	.08	.38
44	7	2D1	.38	.16	.44	.29	.09	.43	.44	.18	.47
45	8	5C1A	.43	.27	.39	.47	.25	.41	.40	.30	.37
46	8	5C2C	.74	.24	.47	.80	.30	.48	.69	.23	.47
47	8	4B1	.56	.29	.49	.58	.28	.47	.55	.32	.50
48	8	4B1	.73	.27	.43	.75	.33	.52	.72	.22	.34
49	8	5C1H	.66	.06	.35	.58	.17	.41	.71	-.07	.31
50	8	4C2	.58	.31	.42	.49	.37	.47	.64	.24	.40

TABLE 4-14

Item Analysis Data - Grade 12 Listening Field Test

			MF = 154			M = 80			F = 74		
ITEM	SCALE	OBJ.	DIFF.	R(TOTAL)	R(SCALE)	DIFF.	R(TOTAL)	R(SCALE)	DIFF.	R(TOTAL)	R(SCALE)
1	1	5C1A	.94	.39	.37	.93	.26	.44	.34	.95	.22
2	1	5C2H	.85	.20	.55	.96	.15	.60	.95	.20	.34
3	1	4C5	.72	.22	.58	.62	.11	.59	.82	.31	.51
4	1	5C2C	.34	.25	.67	.36	.22	.66	.32	.30	.75
5	2	4B3	.77	.19	.29	.79	.08	.33	.76	.33	.25
6	2	4B7	.55	.04	.30	.64	.02	.28	.45	.13	.31
7	2	4C2	.68	.25	.43	.72	.18	.41	.64	.36	.44
8	2	5C2H	.77	.15	.24	.80	.25	.19	.74	.07	.27
9	2	5C10	.33	.15	.32	.30	-.03	.26	.36	.33	.37
10	2	5C1C	.42	.19	.40	.36	.14	.38	.47	.22	.44
11	2	5C2F	.47	.32	.45	.48	.36	.42	.47	.30	.48
12	2	5C3B	.33	.14	.30	.31	.10	.22	.35	.18	.38
13	2	5C10	.75	.16	.27	.75	.23	.30	.74	.09	.25
14	3	5C2H	.79	.17	.29	.79	.15	.26	.80	.18	.33
15	3	5C1P	.75	.05	.19	.75	.03	.07	.74	.08	.31
16	3	5C1G	.68	.06	.35	.68	.07	.41	.69	.05	.30
17	3	5C1G	.60	.30	.38	.51	.16	.38	.70	.42	.37
18	3	4B7	.79	.21	.22	.81	.31	.25	.75	.16	.29
19	3	5C30	.36	.18	.41	.41	.43	.31	.27	.42	
20	3	5C3E	.53	-.06	.33	.58	.00	.34	.49	-.11	.33
21	3	5C2F	.71	.32	.40	.69	.31	.49	.74	.32	.32
22	3	5C2C	.35	.18	.20	.30	.10	.07	.41	.23	.30
23	4	5C1P	.76	.16	.73	.73	.02	.69	.80	.32	.78
24	4	5C1P	.86	.23	.54	.84	.20	.46	.89	.25	.64
25	4	5C1P	.87	.29	.59	.85	.27	.56	.89	.32	.64

TABLE 4-14 (Continued)

Item Analysis Data - Grade 12 Listening Field Test

			MF = 154			M = 80			F = 74		
ITEM	SCALE	OBJ.	DIFF.	R(TOTAL)	R(SCALE)	DIFF.	R(TOTAL)	R(SCALE)	DIFF.	R(TOTAL)	R(SCALE)
26	5	5C3E	.67	.16	.35	.66	.21	.34	.68	.11	.35
27	5	5C1A	.42	.11	.44	.45	.07	.39	.38	.19	.49
28	5	5C1E	.68	.24	.48	.71	.38	.47	.65	.12	.49
29	5	5C1C	.56	.25	.32	.53	.26	.30	.59	.23	.34
30	5	5C1D	.53	.20	.39	.57	.28	.42	.49	.15	.35
31	5	5C3E	.29	.14	.34	.34	.18	.15	.23	.16	.55
32	5	5C1H	.71	.22	.31	.68	.15	.43	.74	.28	.18
33	5	5C1G	.56	.24	.43	.51	.27	.44	.62	.16	.44
34	6	5C2C	.60	.29	.69	.54	.22	.73	.66	.35	.64
35	6	2D1	.61	.20	.55	.48	.18	.67	.76	.13	.30
36	6	5C30	.68	.03	.46	.75	.20	.53	.61	-.09	.47
37	6	5C30	.72	.27	.50	.70	.39	.45	.74	.11	.56
38	7	5C2F	.55	.09	.55	.57	.07	.55	.53	.14	.57
39	7	5C30	.49	.07	.63	.43	.12	.62	.55	-.02	.62
40	7	2B1	.78	.24	.33	.76	.25	.37	.79	.21	.28
41	7	5C3B	.42	.18	.52	.37	.24	.59	.49	.07	.43
42	8	2B1	.72	.27	.51	.75	.32	.61	.69	.24	.38
43	8	2D1	.70	.10	.47	.66	.09	.48	.74	.08	.49
44	8	5C1F	.85	.23	.46	.86	.28	.47	.84	.19	.45
45	8	4C2	.32	.05	.53	.39	.29	.61	.24	-.19	.41
46	9	4B3	.62	.13	.27	.60	.14	.33	.64	.17	.24
47	9	4B7	.65	.27	.53	.58	.28	.51	.72	.21	.51
48	9	4B1	.75	.26	.54	.72	.25	.61	.78	.25	.44
49	9	4B1	.58	.19	.51	.51	.22	.45	.66	.11	.55
50	9	4C5	.27	.11	.45	.24	.11	.37	.30	.09	.53

based on the total test. The latter three statistics are presented first for the total group and then for male and female populations separately.

One of the first problems that faced the Research Team was to determine which of the two available point biserials would be most useful as a measure of discrimination. Had the test been norm-referenced, it would have been appropriate to use the statistic based on the total test. This seemed to us, however, to place too stringent a requirement on a number of items which appear to be valid indications of listening ability. In the analysis of both Pilot and Field Tests, therefore, we decided to use the point biserial coefficients associated with the stimulus sub-scales as criteria for item selection, revision, and deletion. Using this criterion, and a cut-off point of .20 or lower for the total sample as being inadequate for discrimination, it can be seen that Item 29 on the Grade 3 test, Item 31 on the Grade 6 test, Item 41 on the Grade 9 test, and Item 15 on the the Grade 12 test fall short of this criterion.

A perusal of male-female differences indicates slight variations in the discriminating power of items; however, in a sample of this size such variations are to be expected and it is our feeling that, in general, the items are acceptable in the sense that good students are choosing the correct answers and the poorer students are being mislead by incorrect opinions.

In both the Pilot and Field Test analysis it was decided that a Difficulty Index of .80 or less would be accepted as adequate. We note that in Grade 3 Items 13, 20, 22, 25 and 33 have a Difficulty Index above this point. In Grade 6 this is true for Items 7, 10, 26, 33, and 36; in Grade 9 for Items 4, 15, 18, 33 and 35; and in Grade 12 for Items 1, 2, 24, 25 and 44. Male-female differences do not appear to be significant. Revision of these items was considered, but again it should be noted that the criterion for inclusion of an item on the basis of difficulty alone would not be appropriate in a criterion-referenced test. That is, even an easy item may be a valid one if it measures an important objective and if it discriminates among weak and strong students.

Overall, then, these tests appear to meet the criteria of a "good" test in terms of item analysis data. The discriminating power of most items is adequate; the difficulty of individual items with some revision is adequate; and the average level of difficulty seems appropriate.

Speaking (Written)

The same decision rules and criteria were applied to the written part of the Speaking Test. Tables 4-15 to 4-18 provide the item analysis data for the Grades 3, 6, 9 and 12 Written Speaking Tests.

In the Grade 3 test all items appear to discriminate above the .20 level. This is true whether we use the total test or the subscales against which to measure homogeneity of items. For Grade 6 we see that all items meet the established criterion. This is true also of Grades 9 and 12. No significant male-female differences can be seen. Overall, then, the items appear to discriminate and no revisions to meet this criterion are required.

Item difficulty level provided us with more of a problem. In Grade 3 Items 2, 3, 4, 6, 7 and 8 all appear to have been easy for students. This indicates that students at this level are generally able to meet the objectives which the Steering Committee felt were appropriate to this level. While it could be argued that the items are too easy and should be changed, it should be noted that the written portion of the test was given as preparation for the production of speech in later parts of the test. It also acted as a

TABLE 4-15

Item Analysis Data - Grade 3 Speaking

Field Test (Written)

			MF = 214			M = 105			F = 109		
ITEM	SCALE	OBJ.	DIFF.	R(TOTAL)	R(SCALE)	DIFF.	R(TOTAL)	R(SCALE)	DIFF.	R(TOTAL)	R(SCALE
1	1	4C6	.74	.46	.57	.78	.42	.48	.69	.50	.63
2	1	3B6	.84	.33	.46	.82	.21	.39	.85	.45	.54
3	1	3B5	.83	.50	.64	.83	.44	.58	.83	.55	.69
4	1	3B5	.87	.40	.49	.89	.33	.42	.86	.46	.54
5	1	4C6	.62	.49	.57	.66	.49	.55	.58	.49	.59
6	2	4C1	.89	.40	.54	.90	.36	.46	.88	.44	.63
7	2	4C1	.97	.37	.44	.97	.34	.43	.96	.39	.46
8	2	4C5	.85	.36	.51	.81	.39	.56	.89	.34	.45
9	2	4C1	.70	.29	.45	.66	.37	.48	.74	.22	.41
10	2	1D1	.67	.36	.57	.63	.37	.49	.71	.35	.63

TABLE 4-16

Item Analysis Data - Grade 6 Speaking

Field Test (Written)

			MF = 228			M = 109			F = 119		
ITEM	SCALE	OBJ.	DIFF.	R(TOTAL)	R(SCALE)	DIFF.	R(TOTAL)	R(SCALE)	DIFF.	R(TOTAL)	R(SCALE)
1	1	4B5	.57	.38	.38	.59	.40	.40	.56	.36	.37
2	1	3B5	.72	.49	.51	.69	.47	.49	.75	.50	.52
3	1	3B6	.77	.48	.49	.72	.47	.47	.82	.48	.51
4	1	3B6	.44	.27	.36	.38	.25	.36	.50	.29	.35
5	1	1A3	.65	.43	.47	.61	.40	.46	.69	.47	.47
6	1	4C6	.42	.27	.34	.50	.30	.33	.35	.28	.38
7	1	1D1	.60	.43	.49	.60	.46	.53	.59	.41	.45
8	1	1E1	.60	.44	.48	.61	.46	.52	.59	.43	.46
9	1	1A3	.33	.36	.39	.32	.45	.46	.34	.27	.32
10	2	1A4	.92	.34	.40	.94	.32	.35	.91	.36	.44
11	2	1E1	.84	.40	.49	.86	.38	.47	.82	.43	.52
12	2	3B5	.57	.32	.44	.64	.37	.47	.50	.30	.42
13	2	4C6	.59	.32	.47	.58	.30	.48	.62	.34	.47
14	2	1D1	.68	.42	.49	.61	.41	.54	.75	.43	.46
15	2	4B5	.70	.44	.58	.70	.46	.57	.70	.43	.60

TABLE 4-17

Item Analysis Data - Grade 9 Speaking

Field Test (Written)

			MF = 256			M = 113			N = 143		
ITEM	SCALE	OBJ	DIFF.	R(TOTAL)	R(SCALE)	DIFF.	R(TOTAL)	R(SCALE)	DIFF.	R(TOTAL)	R(SCALE)
1	1	4B5	.76	.33	.37	.77	.39	.41	.76	.29	.35
2	1	1A3	.60	.25	.36	.63	.24	.36	.58	.28	.37
3	1	1E1	.81	.36	.35	.83	.32	.35	.80	.41	.54
4	1	1D1	.60	.40	.56	.54	.39	.57	.65	.39	.54
5	1	1D1	.91	.29	.30	.88	.33	.36	.93	.22	.23
6	1	1B1	.62	.20	.34	.61	.22	.29	.64	.18	.38
7	1	1A3	.75	.35	.43	.67	.23	.41	.81	.45	.44
8	1	2B3	.74	.39	.44	.66	.40	.49	.80	.37	.38
9	1	1E1	.73	.19	.31	.72	.20	.25	.74	.20	.35
10	2	4B5	.96	.18	.26	.97	.05	.19	.96	.28	.31
11	2	3C6	.91	.25	.43	.91	.24	.38	.92	.26	.47
12	2	3B5	.89	.21	.32	.83	.23	.33	.94	.13	.33
13	2	3B6	.83	.32	.52	.84	.27	.46	.83	.37	.56
14	2	3B6	.79	.40	.59	.82	.35	.48	.77	.46	.67
15	2	4B5	.97	.19	.32	.95	.20	.43	.99	.16	.22
16	2	4C6	.60	.46	.66	.55	.39	.60	.64	.51	.70
17	3	4C1	.72	.25	.39	.70	.31	.39	.73	.19	.39
18	3	4C4	.61	.52	.51	.56	.51	.56	.65	.51	.45
19	3	4C4	.43	.29	.43	.36	.23	.39	.49	.32	.45
20	3	4C9	.49	.24	.41	.52	.35	.47	.47	.16	.37
21	3	4C9	.62	.54	.64	.54	.60	.68	.68	.48	.59
22	3	4C5	.82	.52	.64	.79	.54	.68	.83	.50	.60
23	3	4C5	.51	.28	.45	.52	.39	.43	.50	.20	.47

TABLE 4-18

Item Analysis Data - Grade 12 Speaking
Field Test (Written)

				MF = 151			M = 81			F = 70	
ITEM	SCALE	OBJ.	DIFF.	R(TOTAL)	R(SCALE)	DIFF.	R(TOTAL)	R(SCALE)	DIFF.	R(TOTAL)	R(SCALE)
1	1	1D1	.91	.54	.53	.91	..24	.29	.90	.76	.72
2	1	1E1	.82	.50	.54	.79	.45	.45	.87	.61	.67
3	1	1A4	.90	.39	.44	.89	.10	.18	.91	.66	.70
4	1	1D1	.82	.39	.44	.84	.19	.28	.81	.55	.57
5	1	4B5	.58	.33	.39	.56	.24	.26	.61	.42	.51
6	1	4B5	.63	.33	.38	.69	.42	.53	.55	.29	.27
7	1	3B6	.55	.39	.49	.51	.34	.48	.60	.46	.53
8	1	3B6	.87	.43	.52	.86	.16	.25	.88	.67	.77
9	1	3B6	.78	.38	.48	.78	.21	.43	.78	.52	.53
10	1	2B3	.68	.41	.41	.60	.33	.30	.76	.55	.56
11	1	1A4	.53	.14	.30	.52	.07	.24	.54	.21	.36
12	2	4C1	.82	.45	.49	.83	.26	.28	.81	.60	.68
13	2	4C4	.78	.44	.52	.79	.22	.32	.78	.65	.74
14	2	4C4	.47	.24	.48	.52	.30	.52	.42	.21	.46
15	2	4C9	.47	.15	.31	.43	.17	.31	.52	.15	.33
16	2	4C9	.72	.28	.23	.69	.22	.23	.75	.35	.24
17	2	4C9	.65	.53	.62	.59	.45	.57	.72	.65	.70
18	2	4C5	.79	.57	.74	.79	.48	.70	.79	.70	.80
19	2	4C5	.67	.39	.43	.67	.17	.31	.67	.56	.55

motivational device which at this grade level seems to us to be of particular importance. Thus, with some minor changes, to be discussed later, we feel it should remain intact, for though the items are easy, they do discriminate well.

In Grade 6 the problem of easy items does not appear. Only Items 10 and 11 fell below criterion. This was not unexpected as at this grade level the items were substantive rather than acting only as preparation. The same comment may be made for Grade 9, where Items 3, 5, 10, 11, 12, 13, 15 and 22, and Grade 12, where Items 1, 2, 3, 4, 8 and 12 appear to have been somwhat easy. Most of these items could be revised by making misleading options more attractive. The basic consideration must, nevertheless, remain the validity with which an item measures an important objective. This validity derives from the judgment of experts rather than from item analysis data.

The Written Speaking Tests thus appear to be generally adequate. The general difficulty appears to be good in terms of motivation; the items discriminate well for both male and female groups; and the difficulty of specific items, after some revision, should produce adequate standards of performance.

Speaking (Ratings)

Tables 4-19 to 4-22 provide the data which describe the rating scales used to evaluate the production of speech. A rating of "1" on this scale is to be interpreted as a high rating. On each table the percentage distribution is provided for each item for the total group and then for male and female populations. This is followed in each case by the mean rating for that item.

In general, we note that there is a good spread of ratings, indicating that raters were able to discriminate among students. We also note that, in general, females were rated slightly higher

TABLE 4-19

Item Analysis Data - Grade 3 Speaking Field Test

(Rating Scales)

		MF = 213						M = 102						F = 111					
ITEM	OBJ.	RATING (%)					$\bar{X}$	RATING (%)					$\bar{X}$	RATING (%)					$\bar{X}$
NO.	NO.	1	2	3	4	5		1	2	3	4	5		1	2	3	4	5	
11	1C2	.01	.20	.59	.19	.01	2.9	.01	.34	.57	.18	.01	2.9	.01	.17	.60	.21	.00	3.0
12	1E3	.00	.51	.39	.08	.02	2.6	.00	.49	.37	.09	.04	2.7	.00	.53	.40	.72	.00	2.5
13	2B1	.01	.09	.21	.46	.23	3.7	.00	.06	.22	.46	.27	3.9	.03	.12	.22	.45	.19	3.7
14	2B2	.00	.11	.43	.42	.03	3.3	.00	.07	.45	.44	.04	3.5	.00	.15	.41	.41	.02	3.3
15	3B2	.01	.33	.48	.16	.01	2.8	.01	.29	.55	.11	.03	2.9	.00	.37	.41	.22	.00	2.9
16	4C5	.14	.65	.16	.04	.01	2.1	.15	.67	.13	.03	.01	2.1	.14	.63	.19	.05	.00	2.1
17	4C2	.09	.55	.24	.03	.01	2.2	.09	.55	.25	.02	.01	2.3	.09	.56	.23	.05	.00	2.3
18	4C4	.09	.50	.31	.09	.01	2.4	.08	.54	.29	.08	.01	2.4	.11	.47	.32	.10	.00	2.4
19	3A1	.10	.44	.30	.15	.01	2.5	.08	.46	.25	.19	.02	2.6	.12	.42	.34	.12	.00	2.5
20	3B2	.00	.14	.45	.35	.05	3.3	.00	.14	.44	.37	.05	3.3	.00	.15	.46	.33	.06	3.3
21	3C8	.14	.56	.26	.02	.01	2.1	.15	.54	.29	.00	.02	2.2	.13	.58	.24	.04	.00	2.2
22	3B1	.12	.60	.23	.02	.02	2.2	.14	.55	.26	.02	.02	2.2	.10	.64	.21	.03	.01	2.2
23	3C7	.11	.50	.27	.04	.01	2.2	.13	.45	.28	.04	.02	2.3	.09	.55	.27	.04	.00	2.3
24	3C8	.14	.30	.37	.15	.03	2.6	.15	.27	.39	.19	.01	2.7	.13	.33	.36	.11	.05	2.6
25	4A1	.07	.57	.27	.07	.02	2.4	.07	.55	.28	.07	.02	2.4	.06	.58	.27	.07	.02	2.4
26	4C5	.16	.61	.20	.02	.00	2.0	.15	.63	.19	.03	.00	2.1	.17	.60	.21	.01	.00	2.1
27	1E2	.02	.32	.50	.13	.03	2.8	.01	.35	.46	.16	.01	2.8	.03	.28	.54	.10	.06	2.9
28	1E4	.00	.02	.08	.42	.27	4.3	.00	.02	.08	.39	.51	4.4	.00	.01	.08	.46	.44	4.3
29	*	.22	.46	.18	.13	.01	2.2	.18	.46	.18	.18	.01	2.4	.26	.46	.18	.07	.00	2.1
30	1A1	.17	.55	.22	.07	.00	2.1	.14	.52	.27	.06	.00	2.3	.10	.57	.16	.07	.00	2.1
31	1C1	.04	.53	.31	.10	.01	2.4	.00	.55	.29	.14	.01	2.6	.08	.51	.32	.06	.00	2.4
32	1C2	.06	.34	.43	.15	.02	2.7	.03	.32	.42	.18	.03	2.9	.09	.35	.44	.09	.01	2.6
33	1D2	.05	.35	.34	.25	.02	2.8	.02	.30	.38	.27	.03	3.0	.07	.34	.31	.23	.01	2.7
34	1E3	.06	.50	.33	.09	.01	2.4	.04	.49	.34	.10	.02	2.6	.07	.51	.32	.09	.00	2.4
35	2B2	.01	.26	.44	.27	.01	3.0	.00	.22	.43	.31	.03	3.2	.03	.30	.44	.23	.00	2.9
36	3A1	.15	.43	.33	.09	.01	2.3	.13	.42	.30	.15	.01	2.5	.16	.44	.35	.04	.01	2.3
37	3A2	.11	.48	.33	.07	.01	2.3	.08	.51	.34	.07	.01	2.4	.14	.45	.32	.06	.00	2.3
38	1E2	.08	.41	.30	.18	.02	2.6	.03	.42	.29	.19	.04	2.8	.12	.40	.31	.17	.01	2.6

TABLE 4-20

Item Analysis Data - Grade 6 Speaking Field Test

(Rating Scales)

		MF = 213						M = 102						F = 111					
ITEM NO.	OBJ. NO.	RATING (%)					$\bar{X}$	RATING (%)					$\bar{X}$	RATING (%)					$\bar{X}$
		1	2	3	4	5		1	2	3	4	5		1	2	3	4	5	
16	1C2	.01	.20	.56	.21	.01	3.0	.01	.21	.53	.23	.00	3.0	.02	.19	.58	.20	.02	3.0
17	1E3	.01	.50	.39	.10	.00	2.6	.01	.44	.43	.12	.00	2.7	.01	.56	.35	.08	.00	2.5
18	2B1	.01	.13	.25	.52	.08	3.5	.02	.12	.22	.56	.08	3.6	.01	.14	.29	.48	.08	3.5
19	2B2	.00	.09	.53	.36	.01	3.3	.00	.12	.52	.34	.00	3.3	.01	.06	.53	.38	.02	3.3
20	3B2	.02	.25	.51	.12	.03	2.9	.03	.23	.51	.19	.03	3.0	.01	.27	.50	.17	.03	2.9
														.03	.27	.51	.14	.02	2.9
21	3C2	.02	.25	.56	.14	.01	2.9	.01	.23	.60	.14	.00	2.9	.03	.27	.51	.14	.02	2.9
22	4C5	.14	.58	.24	.04	.00	2.2	.17	.56	.23	.04	.00	2.1	.11	.61	.24	.04	.01	2.2
23	4C2	.08	.55	.30	.04	.00	2.3	.06	.51	.32	.07	.00	2.4	.09	.59	.28	.01	.01	2.2
24	4C5	.14	.42	.37	.07	.00	2.4	.14	.41	.39	.07	.00	2.4	.15	.43	.35	.07	.01	2.4
25	3A1	.05	.50	.37	.07	.00	2.5	.06	.48	.38	.08	.00	2.5	.04	.53	.35	.06	.01	2.5
26	3B2	.01	.14	.21	.50	.15	3.6	.02	.17	.17	.46	.18	3.6	.01	.11	.25	.52	.12	3.6
27	3C4	.01	.18	.29	.48	.05	3.4	.01	.18	.28	.51	.03	3.4	.01	.17	.30	.45	.07	3.4
28	3C7	.19	.58	.19	.03	.00	2.1	.24	.55	.17	.03	.00	2.0	.14	.60	.22	.03	.00	2.1
29	3C7	.16	.60	.17	.05	.02	2.2	.24	.54	.15	.06	.01	2.1	.09	.62	.21	.05	.03	2.3
30	3C7	.19	.59	.15	.05	.00	2.1	.24	.56	.15	.03	.00	1.9	.13	.63	.17	.06	.00	2.2
31	3C8	.12	.33	.38	.13	.02	2.6	.15	.32	.39	.10	.01	2.5	.10	.34	.38	.16	.03	2.7
32	4A1	.04	.44	.40	.10	.01	2.6	.08	.40	.39	.10	.01	2.6	.00	.47	.42	.09	.02	2.7
33	4C5	.16	.78	.04	.02	.00	1.9	.21	.70	.06	.03	.00	1.9	.11	.86	.03	.01	.00	1.9
34	1E2	.03	.25	.52	.16	.03	2.9	.02	.30	.47	.18	.03	2.9	.04	.21	.59	.14	.03	2.9
35	1E4	.00	.01	.08	.50	.40	4.3	.00	.02	.06	.54	.37	4.3	.00	.00	.11	.46	.44	4.3
36	*	.24	.43	.20	.08	.02	2.2	.22	.40	.23	.13	.03	2.4	.27	.47	.18	.04	.03	2.1
37	1A1	.24	.56	.19	.00	.00	2.0	.29	.52	.19	.00	.00	1.9	.19	.61	.19	.00	.00	2.0
38	1C1	.08	.60	.25	.06	.00	2.3	.04	.61	.29	.05	.00	2.4	.12	.60	.21	.06	.00	2.2
39	1C2	.07	.35	.37	.18	.00	2.7	.05	.39	.35	.19	.00	2.7	.09	.33	.38	.18	.00	2.7
40	1D2	.04	.42	.33	.19	.00	2.7	.03	.45	.29	.22	.00	2.7	.04	.40	.37	.16	.00	2.7

TABLE 4-20 (Continued)

ITEM NO.	OBJ. NO.	MF = 213						M = 102						F = 111					
		RATING (%)						RATING (%)						RATING (%)					
		1	2	3	4	5	$\overline{X}$	1	2	3	4	5	$\overline{X}$	1	2	3	4	5	$\overline{X}$
41	1E3	.05	.52	.33	.09	.00	2.5	.02	.57	.32	.09	.00	2.5	.08	.47	.34	.09	.00	2.5
42	2B2	.03	.29	.56	.10	.00	2.8	.02	.28	.57	.12	.01	2.8	.04	.31	.56	.08	.00	2.7
43	3A1	.10	.57	.27	.04	.00	2.3	.08	.59	.28	.05	.00	2.3	.12	.56	.26	.04	.01	2.2
44	3A2	.12	.51	.29	.05	.01	2.3	.09	.48	.32	.08	.02	2.5	.14	.54	.26	.03	.00	2.2
45	1E2	.05	.48	.32	.12	.00	2.5	.00	.51	.32	.15	.01	2.7	.10	.46	.33	.10	.00	2.4

TABLE 4-21

Item Analysis Data - Grade 9 Speaking Field Test

(Rating Scales)

ITEM NO.	OBJ. NO.	MF = 243 RATING (%) 1	2	3	4	5	$\overline{X}$	M = 108 RATING (%) 1	2	3	4	5	$\overline{X}$	F = 135 RATING (%) 1	2	3	4	5	$\overline{X}$
24	1C1	.07	.49	.35	.09	.00	2.4	.06	.42	.40	.13	.00	2.8	.08	.54	.32	.06	.00	2.4
25	1E3	.05	.57	.28	.08	.00	2.4	.02	.57	.28	.13	.00	2.5	.06	.57	.29	.04	.01	2.3
26	3A1	.14	.58	.24	.04	.00	2.2	.13	.49	.29	.07	.00	2.3	.14	.64	.20	.02	.00	2.1
27	2A2	.03	.18	.39	.31	.08	3.2	.03	.12	.43	.32	.10	3.4	.02	.23	.37	.30	.07	3.2
28	2B2	.06	.16	.14	.38	.26	3.6	.05	.13	.17	.38	.28	3.7	.07	.17	.13	.39	.24	3.6
29	3C2	.04	.54	.35	.05	.00	2.4	.06	.44	.43	.07	.00	2.5	.03	.61	.29	.04	.00	2.4
30	3C8	.15	.56	.21	.07	.00	2.2	.16	.50	.23	.10	.00	2.3	.14	.60	.21	.04	.00	2.1
31	4C2	.07	.63	.22	.05	.00	2.3	.06	.61	.25	.07	.00	2.4	.08	.64	.21	.03	.01	2.2
32	4C4	.09	.33	.43	.14	.00	2.6	.11	.29	.38	.19	.01	2.7	.07	.34	.47	.09	.00	2.6
33	4C4	.09	.31	.24	.21	.10	2.9	.09	.26	.23	.23	.13	3.0	.09	.34	.26	.19	.08	2.8
34	3B3	.11	.61	.24	.03	.00	2.2	.10	.58	.28	.02	.00	2.2	.12	.63	.22	.03	.00	2.2
35	4B3	.14	.33	.39	.09	.02	2.5	.13	.29	.44	.08	.03	2.6	.15	.35	.35	.10	.01	2.4
36	3B2	.04	.15	.46	.32	.03	3.2	.06	.13	.47	.31	.04	3.1	.02	.16	.46	.34	.03	3.2
37	3C7	.38	.51	.09	.02	.00	1.8	.36	.50	.11	.02	.01	1.8	.39	.52	.08	.02	.00	1.7
38	3C7	.43	.51	.07	.00	.00	1.6	.38	.52	.10	.00	.00	1.7	.46	.50	.04	.00	.00	1.6
39	3C4	.01	.15	.49	.33	.02	3.2	.00	.13	.47	.37	.02	3.3	.02	.16	.50	.29	.02	3.1
40	4C5	.22	.51	.21	.05	.00	2.1	.20	.52	.23	.05	.00	2.1	.24	.51	.19	.05	.01	2.1
41	1E2	.07	.41	.32	.18	.02	2.7	.08	.33	.38	.19	.02	2.7	.07	.47	.27	.17	.02	2.6
42	1E4	.01	.09	.33	.51	.05	3.5	.01	.04	.36	.52	.05	3.6	.02	.13	.31	.50	.05	3.5
43	3C5	.00	.01	.01	.06	.90	4.9	.00	.01	.02	.09	.87	4.8	.00	.02	.02	.04	.92	4.9
44	3C5	.00	.01	.02	.03	.93	4.9	.00	.02	.04	.03	.92	4.8	.00	.01	.01	.04	.94	4.9
45	3C5	.00	.01	.02	.01	.95	4.9	.00	.02	.02	.00	.95	4.9	.00	.02	.02	.02	.94	4.9

TABLE 4-21 (Continued)

(Rating Scales)

ITEM NO.	OBJ. NO.	MF = 243 RATING (%) 1	2	3	4	5	$\overline{X}$	M = 108 RATING (%) 1	2	3	4	5	$\overline{X}$	F = 135 RATING (%) 1	2	3	4	5	$\overline{X}$
46	1A1	.21	.04	.14	.01	.00	1.9	.18	.68	.11	.03	.00	2.0	.23	.60	.17	.00	.00	1.9
47	1C1	.15	.73	.12	.00	.00	2.0	.11	.76	.11	.01	.00	2.0	.17	.69	.13	.00	.00	1.9
48	1C2	.14	.33	.40	.12	.00	2.5	.10	.35	.39	.15	.00	2.6	.16	.32	.41	.10	.00	2.5
49	1D2	.10	.58	.26	.05	.00	2.3	.08	.56	.32	.05	.00	2.3	.13	.60	.22	.05	.00	2.2
50	1E5	.08	.25	.49	.16	.00	2.8	.06	.22	.53	.18	.01	2.9	.09	.28	.46	.15	.00	2.7
51	1E3	.07	.39	.39	.11	.00	2.6	.04	.36	.41	.11	.00	2.6	.09	.42	.37	.11	.00	2.5
52	2B2	.01	.03	.22	.35	.39	4.1	.00	.03	.19	.34	.41	4.1	.01	.02	.24	.36	.37	4.0
53	3A1	.19	.52	.25	.03	.00	2.1	.16	.43	.36	.04	.01	2.3	.23	.59	.16	.02	.00	2.0
54	1E2	.18	.57	.18	.05	.01	2.1	.17	.48	.26	.07	.02	2.3	.19	.65	.13	.03	.00	2.0

TABLE 4-22

Item Analysis Data - Grade 12 Speaking Field Test

(Rating Scales)

ITEM NO.	OBJ. NO.	MF = 151 RATING (%) 1	2	3	4	5	$\overline{X}$	M = 80 RATING (%) 1	2	3	4	5	$\overline{X}$	F = 71 RATING (%) 1	2	3	4	5	$\overline{X}$
20	1C1	.22	.42	.32	.02	.01	2.1	.24	.34	.38	.03	.01	2.2	.20	.51	.25	.01	.00	2.0
21	1E3	.20	.48	.26	.02	.00	2.1	.19	.43	.31	.04	.00	2.2	.22	.54	.20	.01	.00	2.0
22	3A1	.29	.51	.13	.03	.00	1.9	.30	.45	.17	.05	.00	1.9	.28	.58	.08	.01	.00	1.8
23	2A2	.11	.28	.40	.16	.01	2.6	.11	.29	.38	.17	.02	2.7	.11	.28	.42	.15	.00	2.6
24	2B2	.14	.26	.36	.12	.09	2.7	.13	.26	.32	.16	.08	2.7	.14	.27	.40	.07	.10	2.7
25	3C2	.24	.38	.30	.04	.00	2.1	.16	.38	.34	.07	.00	2.3	.32	.40	.25	.00	.00	1.9
26	3C8	.27	.45	.23	.02	.00	2.0	.21	.41	.31	.04	.00	2.1	.34	.48	.14	.01	.00	1.8
27	4C2	.22	.43	.26	.04	.01	2.1	.19	.45	.24	.07	.02	2.2	.25	.40	.30	.01	.00	2.0
28	4C4	.20	.26	.38	.11	.01	2.4	.19	.21	.44	.12	.01	2.5	.22	.32	.32	.10	.00	2.3
29	48C4	.15	.28	.32	.13	.04	2.5	.11	.26	.35	.17	.05	2.7	.20	.31	.28	.08	.02	2.3
30	3B3	.28	.47	.18	.02	.01	1.9	.21	.48	.22	.05	.01	2.1	.35	.47	.14	.00	.00	1.7
31	4B3	.16	.49	.26	.03	.02	2.2	.11	.55	.25	.03	.04	2.3	.21	.42	.27	.04	.00	2.1
32	3B2	.29	.47	.18	.02	.01	1.9	.26	.41	.25	.05	.00	2.0	.32	.54	.11	.00	.01	1.8
33	3C7	.22	.50	.23	.02	.01	2.0	.20	.45	.29	.02	.01	2.1	.24	.55	.17	.02	.00	1.9
34	3C7	.29	.47	.18	.02	.01	1.9	.22	.45	.25	.01	.03	2.1	.37	.50	.10	.02	.00	1.7
35	3C4	.21	.42	.29	.05	.01	2.2	.16	.39	.36	.05	.01	2.3	.27	.45	.21	.05	.00	2.0
36	4C3	.18	.46	.28	.04	.01	2.2	.16	.40	.34	.06	.00	2.3	.21	.52	.22	.01	.01	2.0
37	4C5	.11	.38	.38	.07	.01	2.4	.08	.38	.39	.09	.02	2.5	.14	.40	.37	.05	.00	2.3
38	1E2	.20	.43	.27	.06	.01	2.2	.15	.36	.34	.10	.02	2.4	.27	.50	.20	.01	.00	1.9
39	1E4	.15	.32	.30	.16	.03	2.5	.15	.34	.25	.20	.02	2.6	.15	.31	.35	.12	.02	2.5
40	3C5	.10	.43	.26	.08	.01	2.3	.06	.40	.32	.09	.01	2.5	.15	.45	.20	.07	.00	2.2
41	3C5	.06	.16	.39	.26	.02	3.0	.04	.12	.40	.29	.03	3.1	.08	.21	.38	.22	.00	2.8
42	3C5	.06	.10	.29	.34	.06	3.2	.03	.09	.29	.40	.06	3.4	.11	.11	.30	.28	.05	3.0
43		.41	.44	.11	.01	.01	1.7	.48	.36	.11	.01	.01	1.6	.34	.52	.11	.00	.00	1.7
44		.23	.48	.24	.02	.00	2.0	.22	.41	.31	.01	.00	2.1	.23	.55	.16	.02	.00	1.9

TABLE 4-22 (Continued)

ITEM NO.	OBJ. NO.	MF = 151 RATING (%) 1	2	3	4	5	$\overline{X}$	M = 80 RATING (%) 1	2	3	4	5	$\overline{X}$	F = 71 RATING (%) 1	2	3	4	5	$\overline{X}$
45		.10	.38	.37	.09	.02	2.5	.10	.34	.40	.10	.02	2.6	.11	.43	.34	.08	.01	2.4
46		.13	.32	.38	.11	.03	2.5	.09	.38	.30	.15	.06	2.7	.18	.26	.47	.06	.00	2.4
47		.13	.27	.43	.13	.01	2.5	.11	.27	.40	.17	.01	2.6	.15	.27	.45	.08	.00	2.4
48		.18	.47	.27	.05	.01	2.2	.19	.41	.31	.05	.01	2.2	.18	.52	.22	.04	.01	2.1
49		.02	.11	.40	.34	.09	3.3	.02	.12	.40	.32	.09	3.3	.01	.11	.40	.36	.09	3.4
50		.30	.44	.15	.07	.01	2.0	.32	.36	.19	.10	.00	2.0	.27	.52	.12	.04	.01	1.9
51		.29	.27	.27	.12	.02	2.2	.34	.20	.26	.15	.02	2.3	.25	.34	.27	.08	.01	2.2

265

than males. Whether this represents a legitimate difference in ability or the operation of rater bias is not clear at this time. Nevertheless, we feel that the rating scales are adequate and valid measures of the skills they are designed to measure.

DESCRIPTIVE STATISTICS

Listening

Tables 4-23 to 4-26 present the descriptive statistics for the Listening Field Tests. We have included the means and standard deviations for each sub-test and for male and female groups. The column headed "Alphas" can be taken as a reliability estimate, but it should be noted that where the variance (sigma) falls below 1.000, no reliability estimate is available.

Perusal of Tables 4-23 to 4-26 indicates that the reliabilities of the total test ranged from .43 in Grade 12 to .67 in Grade 6 for the total group. These are not high reliabilities in terms of what one usually expects on a standardized test, but we must recall that reliability estimates of this type are usually based on larger samples, and, more importantly, they are usually provided for norm-referenced tests in which the items are more homogeneous by design. A more appropriate criterion to be applied here would be that of test- re- test reliability which this study was not able to provide. That is, given that objectives have been appropriately set and that adequate items have been devised, a measure of stability is probably of more significance than a measure of homogeneity.

In terms of reliability some interesting anomalies appear. In Grade 3 we note, for example, that the reliability for males on Stimulus Sub-scale 2 is considerably lower than that for females. The reverse is true for Stimulus Sub-scale 5. Since Stimulus Sub-scale 2 involved questions about a recipe, it is tempting to conclude that males were not as well tuned to this stimulus. Sub-scale 5 was based on conversation, in which the measurement appears to be more reliable for boys. Both of these results may, of course,

TABLE 4-23

Statistics for Stimulus Sub-scales and Total Grade 3
Listening Field Test MF=211; M=99; F=112

SCALE	NO. OF ITEMS		MEANS	SIGMAS	ALPHAS
1	4	MF	2.711	.967	0.0000
		M	2.697	.870	0.0000
		F	2.723	1.045	.3007
2	6	MF	4.057	1.309	.3016
		M	3.889	1.238	.1409
		F	4.205	1.351	.3941
3	5	MF	3.199	.958	0.0000
		M	3.192	1.002	.0450
		F	3.205	.917	0.0000
4	14	MF	9.773	2.152	.4623
		M	9.717	2.099	.4163
		F	9.821	2.197	.4966
5	6	MF	3.052	1.263	.2826
		M	2.990	1.352	.3780
		F	3.107	1.175	.1608
TOTAL	35	MF	22.791	4.080	.6005
		M	22.485	3.935	.5585
		F	23.063	4.186	.6280

TABLE 4-24

Statistics for Stimulus Sub-scales and Total Grade 6

Listening Field Test MF=225; M=108; F=117

SCALE	NO. OF ITEMS		MEANS	SIGMAS	ALPHAS
1	5	MF	2.867	1.062	.0050
		M	2.806	.995	0.0000
		F	2.898	1.130	.1159
2	6	MF	4.498	1.136	.1816
		M	4.509	1.101	.1393
		F	4.390	1.296	.3811
3	5	MF	3.133	1.120	.1575
		M	3.009	1.182	.2326
		F	3.212	1.064	.0843
4	4	MF	1.916	1.031	.1300
		M	1.759	1.070	.2024
		F	2.051	.973	0.0000
5	6	MF	3.929	1.403	.4157
		M	3.796	1.373	.3789
		F	4.000	1.456	.4708
6	10	MF	5.827	1.731	.2958
		M	5.843	1.744	.3114
		F	5.788	1.717	.2799
7	4	MF	2.356	1.040	.1570
		M	2.389	.961	0.0000
		F	2.314	1.102	.2890
TOTAL	40	MF	24.524	4.881	.6504
		M	24.111	5.023	.6727
		F	24.653	4.989	.6655

TABLE 4-25

Statistics for Stimulus Sub-scales and Total Grade 9

Listening Field Test MF=244; M=108; F=136

SCALE	NO. OF ITEMS		MEANS	SIGMAS	ALPHAS
1	4	MF	2.607	1.005	.2836
		M	2.594	1.017	.2814
		F	2.619	.991	0.0000
2	9	MF	4.656	1.651	.2600
		M	4.660	1.636	.2605
		F	4.664	1.657	.2610
3	9	MF	5.865	1.461	.1753
		M	5.623	1.581	.2880
		F	6.052	1.334	.0281
4	3	MF	1.471	.797	0.0000
		M	1.396	.843	0.0000
		F	1.522	.760	0.0000
5	8	MF	5.143	1.343	.1557
		M	5.038	1.473	.3060
		F	5.231	1.227	-.0064
6	4	MF	2.557	1.000	.2344
		M	2.358	1.011	.2361
		F	2.701	.970	0.0000
7	7	MF	4.037	1.206	.0585
		M	3.943	1.250	.1395
		F	4.119	1.166	.0019
8	6	MF	3.697	1.207	.0884
		M	3.670	1.294	.2431
		F	3.709	1.132	-.0535
TOTAL	50	MF	30.033	4.887	.5743
		M	29.283	5.266	.6325
		F	30.619	4.507	.5076

TABLE 4-26

Statistics for Stimulus Sub-scales and Total Grade 12
Listening Field Test MF=154; M=80; F=74

SCALE	NO. OF ITEMS		MEANS	SIGMAS	ALPHAS
1	4	MF	2.851	.874	0.0000
		M	2.675	.959	0.0000
		F	3.041	.725	0.0000
2	9	MF	5.058	1.415	.0349
		M	5.138	1.292	.1238
		F	4.973	1.533	.1667
3	9	MF	5.688	1.277	.1265
		M	5.613	1.220	.2482
		F	5.770	1.331	.0005
4	3	MF	2.494	.705	0.0000
		M	2.413	.685	0.0000
		F	2.581	.717	0.0000
5	8	MF	4.403	1.479	.1859
		M	4.425	1.430	.1023
		F	4.378	1.531	.2766
6	4	MF	2.695	1.015	.2412
		M	2.538	1.106	.4090
		F	2.865	.875	0.0000
7	4	MF	1.571	.952	0.0000
		M	1.425	.959	0.0000
		F	1.730	.920	0.0000
8	4	MF	2.571	.859	0.0000
		M	2.635	.954	0.0000
		F	2.514	.740	0.0000
9	5	MF	2.312	1.010	.1139
		M	2.125	1.005	.0515
		F	2.514	.976	0.0000
TOTAL	50	MF	29.643	4.119	.4329
		M	28.975	4.189	.4898
		F	30.365	3.917	.3807

be artifacts of the data and should not be overinterpreted. Similar discrepancies appear from time to time in other grades. It is our feeling that such discrepancies are to be expected and should not be over-interpreted until a larger sample is used. They are also artifactual to some extent because they are based on stimulus sub-scales rather than on logical domains and are often found for sub-scales which contain too few items to be meaningfully interpreted.

With reference to means we note that with the exception of Stimulus Sub-scales 2, 6 and 7 in Grade 6 and Stimulus Sub-scales 2, 5 and 8 in Grade 12, females score slightly lower on all tests than males. The difference is small and not unexpected. In general, there is therefore no reason to believe that the tests are biased in terms of male-female differences. It is possible, of course, that with larger, more heterogeneous samples differences may appear, but given the relative similarities among these groups no further breakdown of data appeared to be necessary.

Speaking (Written)

Tables 4-27 to 4-30 present similar data for the written part of the Speaking Test. Reliabilities range from .39 to .64 for Grade 3 and Grade 12. Since these are relatively short tests composed of heterogeneous items, this reliability estimate is about what one would expect. It is difficult to explain, however, why there are such large differences in the reliability estimates between male and female students. It is possible that male students were not motivated by the stimuli and therefore did not attend consistently to the items.

Male-female differences in achievement level are of about the same magnitude as those in the Listening Tests and are in the expected direction. Differences related to sex in terms of reliability can also be noted. For example, in the Grade 9 test fairly substantial differences are noted in Stimulus Sub-scales 2 and 3.

TABLE 4-27

Statistics for Stimulus Sub-scales and Total Grade 3
Speaking Field Test (Written) MF=214; M=105; F=109

SCALE	NO. OF ITEMS		MEANS	SIGMAS	ALPHAS
1	5	MF	3.888	1.109	.4193
		M	3.971	.971	0.0000
		F	3.807	1.223	.5485
2	5	MF	4.065	.915	0.0000
		M	3.943	.903	0.0000
		F	4.183	.910	0.0000
TOTAL	10	MF	7.953	1.522	.3875
		M	7.914	1.448	.3039
		F	7.991	1.588	.4612

TABLE 4-28

Statistics for Stimulus Sub-scales and Total Grade 6
Speaking Field Test (Written) MF=228; M=109; F=119

SCALE	NO. OF ITEMS		MEANS	SIGMAS	ALPHAS
1	9	MF	5.092	1.872	.4687
		M	5.000	1.939	.5030
		F	5.176	1.804	.4387
2	6	MF	4.303	1.239	.3292
		M	4.330	1.242	.3485
		F	4.277	1.236	.3283
TOTAL	15	MF	9.395	2.664	.5946
		M	9.330	2.730	.6150
		F	9.454	2.602	.5814

TABLE 4-29

Statistics for Stimulus Sub-scales and Total Grade 9
Speaking Field Test (Written) MF=256; M=113; F=143

SCALE	NO. OF ITEMS		MEANS	SIGMAS	ALPHAS
1	9	MF	6.525	1.568	.3439
		M	6.310	1.575	.3133
		F	6.706	1.541	.3595
2	7	MF	5.969	1.087	.4002
		M	5.876	1.023	.2483
		F	6.042	1.134	.5039
3	7	MF	4.105	1.603	.4538
		M	3.929	1.692	.5155
		F	4.224	1.503	.3774
TOTAL	23	MF	16.599	3.140	.6156
		M	16.115	3.244	.6273
		F	16.972	3.010	.5986

TABLE 4-30

Statistics for Stimulus Sub-scales and Total Grade 12
Speaking Field Test (Written) MF=151; M=80; F=71

SCALE	NO. OF ITEMS		MEANS	SIGMAS	ALPHAS
1	11	MF	7.894	1.967	.5472
		M	7.951	1.594	.2483
		F	7.829	2.324	.7163
2	8	MF	5.285	1.661	.4753
		M	5.272	1.474	.2761
		F	5.300	1.854	.6289
TOTAL	19	MF	13.179	2.999	.6381
		M	13.222	2.228	.2866
		F	13.129	3.695	.7904

CORRELATIONAL ANALYSIS BY STIMULUS SUB-SCALES

In order to look at the relationships among the Stimulus Subscales a correlation matrix was constructed for the Listening and Speaking Tests by grade and for male and female populations. Since sub-scales were heterogeneous in terms of items and objectives being measured, and since they varied in length, these data are suggestive only.

Listening

Tables 4-31 to 4-34 present the correlations among sub-scales for Grades 3, 6, 9 and 12. In looking at Table 4-31 we note that in Grade 3 Sub-scale 4 appears to correlate most strongly with the other sub-scales and the total. This is true for both male and female populations. This result probably came about because this was the longest sub-scale in the test and no great significance should be attached to it without further data. The finding that the stimulus sub-scale based on poetry correlated .30 for boys and .50 for girls may, however, be significant though no feasible interpretation can be given at this time.

Table 4-32 presents similar results for the Grade 6 sample. With the exception of Sub-scale 1 the highest correlation for all sub-scales was with questions based on Student Talk relative to the Student Council discussion. Again this was the longest sub-scale and therefore contributed most to the total test. Male-female differences appear to be small and insignificant.

Table 4-33 presents the correlation matrix for the Grade 9 Listening Test. Small, and for the most part, insignificant

Table 4-31

Correlational Analysis of Grade 3 Listening Field Test by Stimulus Sub-scale and Sex

MF = 211

Sub-scale	1	2	3	4	5	Total
1	1.00	.08	.12	.24	-.02	.42
2	.08	1.00	.04	.33	.09	.55
3	.12	.04	1.00	.35	.16	.51
4	.24	.33	.35	1.00	.25	.85
5	-.01	.09	.16	.25	1.00	.50
Total	.41	.55	.51	.85	.50	1.00

M = 99

	1	2	3	4	5	Total
1	1.00	.01	.07	.14	-.13	.30
2	.10	1.00	.06	.28	.04	.51
3	.07-	.06	1.00	.41	.18	.57
4	.14	.28	.41	1.00	.27	.85
5	-.13	.04	.18	.27	1.00	.52
Total	.30	.51	.57	.85	.52	1.00

F = 112

	1	2	3	4	5	Total
1	1.00	.07	.17	.32	.08	.50
2	.07	1.00	.02	.38	.13	.58
3	.17	.02	1.00	.29	.13	.46
4	.32	.38	.29	1.00	.24	.86
5	.08	.13	.13	.24	1.00	.49
Total	.50	.58	.46	.86	.49	1.00

Table 4-32

Correlational Analysis of Grade 6 Listening Field

Test by Stimulus Sub-scale and Sex

MF = 225

Sub-scale	1	2	3	4	5	6	7	Total
1	1.00	.22	.13	.02	.21	.15	.14	.45
2	.22	1.00	.04	.11	.30	.31	.25	.57
3	.13	.04	1.00	.28	.10	.29	.09	.47
4	.02	.11	.28	1.00	.15	.23	.17	.46
5	.21	.30	.10	.15	1.00	.35	.20	.63
6	.15	.32	.29	.23	.35	1.00	.37	.75
7	.14	.25	.09	.17	.20	.37	1.00	.55
Total	.45	.57	.47	.46	.63	.75	.55	1.00

M = 108

	1	2	3	4	5	6	7	Total
1	1.00	.17	.19	.06	.20	.24	.21	.47
2	.17	1.00	.06	.11	.42	.31	.27	.56
3	.19	.06	1.00	.31	.07	.32	.13	.49
4	.06	.11	.31	1.00	.13	.27	.23	.49
5	.20	.42	.07	.13	1.00	.39	.27	.64
6	.24	.31	.32	.27	.39	1.00	.43	.78
7	.21	.27	.13	.23	.27	.43	1.00	.59
Total	.47	.56	.51	.49	.64	.78	.59	1.00

F = 117

	1	2	3	4	5	6	7	Total
1	1.00	.32	.12	.00	.20	.08	.10	.44
2	.32	1.00	.14	.18	.28	.32	.21	.64
3	.12	.14	1.00	.23	.19	.28	.06	.49
4	+.00	.18	.23	1.00	.22	.22	.14	.46
5	.20	.28	.19	.22	1.00	.31	.16	.64
6	.08	.32	.28	.22	.31	1.00	.34	.71
7	.10	.21	.06	.14	.16	.34	1.00	.50
Total	.44	.64	.49	.46	.64	.71	.50	1.00

Table 4-33

Correlational Analysis of Grade 9 Listening Field

Test by Stimulus Sub-scale and Sex

MF = 244

Sub-scale	1	2	3	4	5	6	7	8	Total
1	1.00	.18	.14	.07	.21	.19	.05	.14	.46
2	.18	1.00	.12	.06	.26	.22	.10	.20	.61
3	.14	.12	1.00	.12	.15	.03	.03	.16	.48
4	.07	.06	.12	1.00	.06	.14	.07	.08	.32
5	.21	.26	.15	.06	1.00	.21	.07	.33	.60
6	.19	.22	.03	.14	.21	1.00	.16	.19	.50
7	.05	.10	.03	.07	.07	.16	1.00	.13	.40
8	.14	.20	.16	.08	.33	.19	.13	1.00	.56
Total	.46	.61	.48	.32	.60	.50	.40	.56	1.00

M = 108

	1	2	3	4	5	6	7	8	Total
1	1.00	.29	.19	.09	.31	.22	.04	.21	.54
2	.29	1.00	.17	.08	.24	.29	.05	.15	.60
3	.19	.17	1.00	.03	.19	-.08	.02	.23	.49
4	.09	.08	.03	1.00	.03	.10	.08	.14	.29
5	.31	.24	.19	.03	1.00	.23	.07	.30	.61
6	.22	.29	-.08	.10	.23	1.00	.22	.24	.49
7	.04	.05	.02	.08	.07	.22	1.00	.26	.41
8	.21	.15	.23	.14	.30	.24	.26	1.00	.62
Total	.54	.60	.49	.29	.61	.49	.41	.62	1.00

F = 136

	1	2	3	4	5	6	7	8	Total
1	1.00	.08	.11	.05	.12	.18	.04	.11	.40
2	.08	1.00	.11	.04	.29	.18	.11	.25	.63
3	.11	.11	1.00	.19	.08	.08	.04	.09	.47
4	.05	.04	.19	1.00	.08	.15	.03	.02	.32
5	.12	.29	.08	.08	1.00	.16	.05	.36	.58
6	.18	.18	.08	.15	.16	1.00	.10	.14	.48
7	.04	.11	.04	.03	.05	.10	1.00	.01	.37
8	.11	.25	.09	.02	.36	.14	.01	1.00	.53
Total	.40	.63	.47	.32	.58	.48	.37	.53	1.00

Table 4-34

Correlational Analysis of Grade 12 Listening Field Test by Stimulus Sub-scale and Sex

MF = 154

Sub-scale	1	2	3	4	5	6	7	8	9	Total
1	1.00	.10	.11	.15	.09	.13	-.00	.14	.17	.44
2	.11	1.00	.11	.15	.17	-.01	.01	.09	.08	.53
3	.11	.11	1.00	.14	.11	.10	.00	.12	.02	.49
4	.15	.15	.14	1.00	.08	.13	-.00	-.11	.07	.35
5	.09	.17	.11	.08	1.00	.10	-.06	.03	.04	.51
6	.13	-.01	.10	.13	.10	1.00	-.02	-.05	.04	.35
7	-.00	.01	.00	-.00	-.06	-.02	1.00	.06	.20	.26
8	.14	.09	.12	-.11	.03	-.05	.06	1.00	.07	.32
9	.17	.08	.02	.07	.04	.04	.20	.07	1.00	.42
Total	.44	.53	.49	.35	.51	.35	.26	.32	.42	1.00

M = 80

	1	2	3	4	5	6	7	8	9	Total
1	1.00	.11	-.04	.07	.03	.03	-.04	.16	.21	.36
2	.11	1.00	.07	.04	.25	-.03	-.03	.15	.01	.47
3	-.04	.07	1.00	-.06	.17	.21	.02	.16	.01	.45
4	.07	.04	-.06	1.00	.11	.17	.05	-.12	.10	.27
5	.03	.25	.17	.11	1.00	.08	.04	.25	.11	.61
6	.03	-.03	.21	.17	.08	1.00	-.00	-.09	.18	.40
7	-.04	-.03	.02	.05	.04	-.00	1.00	.10	.12	.29
8	.16	.15	.16	-.12	.25	-.09	.10	1.00	.08	.44
9	.21	.01	.01	.10	.11	.18	.12	.08	1.00	.45
Total	.36	.47	.45	.27	.61	.40	.29	.44	.45	1.00

F = 74

	1	2	3	4	5	6	7	8	9	Total
1	1.00	.13	.29	.21	.19	.22	-.04	.16	.04	.53
2	.13	1.00	.15	.26	.10	.03	.09	.02	.18	.63
3	.29	.15	1.00	.32	.05	-.05	-.04	.09	.00	.53
4	.21	.26	.32	1.00	.05	.06	-.11	-.07	-.00	.43
5	.19	.10	.05	.05	1.00	.14	-.17	-.25	-.01	.43
6	.22	.03	-.05	.06	.14	1.00	-.12	.04	-.23	.25
7	-.04	.09	-.04	-.11	-.17	-.12	1.00	.02	.23	.19
8	.16	.02	.09	-.07	-.25	.04	.02	1.00	.08	.18
9	.04	.18	.00	-.00	-.01	-.23	.23	.08	1.00	.34
Total	.53	.63	.53	.43	.43	.25	.19	.18	.34	1.00

correlations were found. Again Sub-scale 2, the longest of the sub-scales, correlates most with the total. For reasons which are not clear,the Poetry Sub-scale correlates with the total more highly for males than for females. No other major differences appear.

Table 4-34 presents the findings for the Grade 12 sample. Similar results can be observed.

Speaking (Written)

Table 4-35 presents the correlation matrix for the Grade 3 Written Speaking Test. In all cases the correlation between stimulus sub-scales is only moderate. No significant male-female differences appear.

Table 4-36 presents correlations among sub-scales for the Grade 6 Written Speaking Test. Fairly substantial correlations are found suggesting that the skills required for these tests are similar to a degree.

Table 4-37 presents correlational information for the Grade 9 stimulus sub-scales on the written portion of the Speaking Test. It can be seen from Table 4-37 that all three sub-scales correlated about equally with the total. No significant differences in male-female relationships were noted except that there may be a slightly stronger relationship between Sub-scale 3 and the total score for males.

Table 4-38 presents correlational results for Grade 12. In general, we note that as we go up in grade level, the abilities required for the written test become more homogeneous. The fact that for Grade 12 boys the correlation between Sub-scales 1 and 2 is only .05 is difficult to explain. The fact that Sub-scale 1 was based on a poem is highly suggestive and needs further study.

As has already been noted above, the correlations among stimulus sub-scales are likely to be spurious and any comments made based on

Table 4-35

Correlational Analysis of Grade 3 Speaking Field Test (Written) by Stimulus Sub-scales and Sex

MF = 214

Sub-scale	1	2	Total
1	1.00	.12	.80
2	.12	1.00	.69
Total	.80	.69	1.00

M = 105

	1	2	Total
1	1.00	.19	.79
2	.19	1.00	.75
Total	.79	.75	1.00

F = 109

	1	2	Total
1	1.00	.09	.82
2	.09	1.00	.64
Total	.82	.64	1.00

Table 4-36

Correlational Analysis of Grade 6 Speaking Field

Test (Written) by Stimulus Sub-scales and Sex

MF = 228

Sub-scale	1	2	Total
1	1.00	.44	.91
2	.44	1.00	.78
Total	.91	.78	1.00

M = 109

	1	2	Total
1	1.00	.45	.91
2	.45	1.00	.77
Total	.91	.78	1.00

F = 119

	1	2	Total
1	1.00	.45	.91
2	.45	1.00	.78
Total	.91	.78	1.00

Table 4-37

Correlational Analysis of Grade 9 Speaking Field Test (Written) by Stimulus Sub-scales and Sex

MF = 257

Sub-scale	1	2	3	Total
1	1.00	.33	.31	.77
2	.33	1.00	.28	.66
3	.31	.28	1.00	.76
Total	.77	.66	.76	1.00

M = 113

	1	2	3	Total
1	1.00	.27	.38	.77
2	.27	1.00	.34	.63
3	.38	.34	1.00	.81
Total	.77	.63	.81	1.00

F = 143

	1	2	3	Total
1	1.00	.36	.23	.76
2	.36	1.00	.22	.67
3	.23	.22	1.00	.70
Total	.76	.67	.70	1.00

Table 4-38

Correlational Analysis of Grade 12 Speaking Field

Test (Written) by Stimulus Sub-scales and Sex

MF = 151

Sub-scale	1	2	Total
1	1.00	.36	.78
2	.36	1.00	.79
Total	.85	.79	1.00

M = 81

	1	2	Total
1	1.00	.05	.75
2	.05	1.00	.70
Total	.75	.70	1.00

F = 70

	1	2	Total
1	1.00	.55	.90
2	.55	1.00	.85
Total	.90	.85	1.00

them should be treated as suggestive only. Of greater interest to the present study is the determination of the extent to which specific skills are correlated and the extent to which there is correlation between or among speaking skills, listening skills and attitudes. Since the initial sub-scale structure was largely an artifact of the test design, based on the extent to which stimuli lent themselves to item construction, an alternative means of dividing the test into sub-scales was needed. Since the original approach was designed to allocate items to fit a-priori domains, it was decided to use these domains as the basis for a new set of sub-scales which could be entered into a correlation matrix. The results of this analysis are presented on the next page. The problem that faced us, however, was that some of the domains contained very few items and we have thus grouped some domains together where it made sense to do so.

GENERAL CORRELATIONAL ANALYSIS

In order to determine the extent to which abilities in the areas of speaking and listening were related, a Master Correlation Matrix was constructed for each grade. The results of the Attitude Test were also included as a sub-scale. The problem here was to determine which items were to be included in each sub-scale. A previous analysis (See Tables 4-23 to 4-30) already indicated the intercorrelations among sub-scales within each test, but these results, because they were based on stimulus sub-scales, could not be interpreted as correlations among specific abilities or areas of competence. For this purpose it seemed appropriate to deal with the a-priori domains and to treat them as sub-scales. Thus Tables 4-39 to 4-42 summarize the correlations among the domains for speaking, listening and attitude for Grades 3, 6, 9 and 12 respectively. It should be noted that the extent of correlation between any two domains is a reflection of true relationship only to the extent that items were placed in the appropriate domain in the first place. In some cases there were too few items in a domain and these had to be moved into the next most appropriate one. Thus, if items are moved from one domain to another, these relationships would also change. Perhaps of greatest significance are the correlations between speaking, listening, attitude and between written and oral speech. These and other significant differences will be discussed below.

Grade 3

Table 4-39 presents the correlation among domain sub-scales and total scores for speaking, listening and attitude. The table

Table 4-39

Correlation Matrix for Listening, Speaking and Attitude by Domain Sub-scales and Total Test - Grade 3

		L_1	L_2	L_4	L_5	L_6	L_T	S_1	S_2	S_3	S_4	S_5	S_8	SWT	[illegible]	S_T	ATT
L_1	Voice Production	1.00	.34	.25	.13	.14	.65	-.09	.02	-.15	.04	.00	.08	.22	-.07	.04	.00
L_2	Linguistic Factors - Words		1.00	.28	.25	.18	.64	-.29	-.08	-.28	-.01	-.14	-.01	.30	-.26	.22	-.03
L_4	Literal Comprehension			1.00	.13	.28	.69	-.27	-.04	-.17	.07	-.11	-.02	.31	-.21	-.18	-.10
L_5	Interpretative Comprehension				1.00	.33	.53	-.16	-.09	-.15	.04	-.01	.04	.18	-.11	-.09	-.01
L_6	Critical Comprehension					1.00	.54	-.19	-.09	-.15	-.01	-.05	.06	.25	-.15	-.12	.16
L_T	Listening Total						1.00	-.32	-.08	-.29	.01	-.10	.04	.41	-.26	-.21	-.02
S_1	Voice Production							1.00	.54	.70	.35	.49	.42	-.18	.90	.90	.12
S_2	Nonverbal								1.00	.20	.22	.10	.25	.00	.53	.53	.01
S_3	Linguistic Factors - Sounds									1.00	.19	.36	.31	-.17	.68	.67	.07
S_4	Linguistic Factors - Words										1.00	.43	.34	.23	.50	.54	-.03
S_5	Linguistic Factors[1] - Sentences											1.00	.51	-.16	.74	.73	.02
S_8	Organization of Details												1.00	.28	.62	.67	-.01
SWT	Speaking Written													1.00	-.19	-.06	-.14
R	Ratings														1.00	.99	.08
ST	Speaking Total															1.00	.07
ATT	Attitude																1.00

---

1. Contains one item from Domain VI.

indicates generally low relationships among the sub-scales in the Listening Test, the highest being a correlation of .34 between Voice Production and Linguistic Factors - Words. Substantial correlations between the sub-scales and the total test, ranging from .53 to .69, are found as expected. All correlations between Listening sub-scales and the Speaking Total are low and negative, suggesting that the abilities required for these two activities do not overlap. There is, however, a low positive relationship between Speaking (Written) and the Listening sub-scales. Finally, we note that there is very little relationship between Listening and the Attitude measure.

The correlations for the Speaking Test reveal a somewhat different pattern. Correlations among sub-scales range from .10 to .70 and are all positive, indicating that there is some overlap in abilities measured by the sub-scales. It should be noted, however, that some of the commonality is accounted for by the fact that the same rater judged each student on all variables. The sub-scales correlate highly with the total test, ranging from .53 to .90. There is very low correlation between Speaking (Written) and the Oral Speaking sub-scales. This is a bit surprising because one might assume that there is a relationship between knowing how to speak and actually doing it. The correlation between Listening and Speaking Total ($r = -.21$) indicates little relationship between the two. In fact, to the extent that the two tests are valid, it suggests that good speakers are poor listeners. This result, though it may fit our observations of human nature, requires further study. Finally, no relationship appears to exist between the Attitude measure and speaking ability as measured by this test.

Grade 6

Table 4-40 presents the correlations among domain sub-scales

Table 4-40

Correlation Matrix for Listening, Speaking and Attitude by Domain Sub-scales and Total Test - Grade 6

		L_1	L_2	L_4	L_5	L_6	L_T	S_1	S_2	S_3	S_4	S_5	S_8	SWT	R	S_T	ATT
L_1	Voice Production	1.00	.28	.31	.21	.32	.65	-.05	-.10	-.15	.04	.01	.11	.26	-.07	-.02	-.02
L_2	Linguistic Factors - Words		1.00	.26	.25	.16	.60	-.11	-.13	-.17	-.05	-.02	.12	.37	-.12	-.06	-.02
L_4	Literal Comprehension			1.00	.31	.17	.75	-.22	-.15	-.26	.02	-.09	-.01	.38	-.24	-.19	-.20
L_5	Interpretative Comprehension				1.00	.18	.60	-.18	.06	-.21	.01	-.03	.08	.24	-.16	-.13	-.15
L_6	Critical Comprehension					1.00	.52	.04	.07	-.14	-.01	-.04	.03	.26	-.04	.00	-.11
L_T	Listening Total						1.00	-.19	-.10	-.31	.04	-.06	.09	.49	-.22	-.14	-.17
S_1	Voice Production							1.00	.34	.09	.14	.45	.32	-.11	.89	.90	.14
S_2	Nonverbal								1.00	.18	.11	.20	.19	-.14	.34	.33	.08
S_3	Linguistic Factors - Sounds									1.00	.05	.35	.28	-.38	.74	.74	.10
S_4	Linguistic Factors - Words										1.00	.30	.16	.27	.35	.38	.03
S_5	Linguistic Factors - Sentences											1.00	.50	-.17	.60	.58	.04
S_8	Organization of Details 1.												1.00	.08	.59	.62	-.05
SWT	Speaking Written													1.00	-.18	-.07	-.09
R	Ratings														1.00	.99	.10
S_T	Speaking Total															1.00	.09
ATT	Attitude																1.00

1. Contains one item from Domain VI and two items from Domain VII.

and total scores for Speaking, Listening, and Attitudes. We note first of all that in Listening the sub-scales correlate only moderately, .16 to .32. Rather distinct abilities appear to be involved in the sub-scales. The sub-scales are correlated with the total test about equally well, .52 to .75. Again there is no correlation between Speaking (Written) and Listening and a correlation of -.14 between Speaking Total and Listening Total. It appears that at the Grade 6 level, then, there is little if any relationship between Speaking and Listening as measured by these tests. We also note that no relationship exists between Attitude and any of the Listening sub-scales.

On the Speaking Test we note that, as in Grade 3, there are several substantial correlations among sub-scales. For example, Voice Production Factors and Linguistic Factors - Sounds correlate at .69. The sub-scales also correlate highly with the total score. As already pointed out, there is no relationship between Speaking and Listening, nor is there a relationship between Speaking and Attitude measures.

Grade 9

As indicated in Table 4-41, the pattern of correlations among Grade 9 sub-scales is similar to those in Grades 3 and 6. Intercorrelations among Listening sub-scales are low, -.05 to .26. A correlation of .42 between Speaking (Written) and Listening Total suggests again that the ability required to perform the written items is similar to the kind of comprehension ability needed to do well on the Listening items. The very high correlation (.97) between the Rating Scales and the Voice Production sub-scale in Speaking is of interest. A substantial part of this relationship arises from the fact that the first sub-scale had a large number of items and thus formed a major part of the Rating Total. Nevertheless, it appears that a great deal of the variance associated with rating

Table 4-41

Correlation Matrix for Listening, Speaking and Attitude by Domain Sub-scales and Total Test - Grade 9

	L_1	L_2	L	L_4	L_5	L_6	L_T	S_1	S_2	S_4	S_5	S_8	SWT	R	S_T	ATT
L_1 Voice Production	1.00	.07	.05	.03	-.05	.01	.23	.03	-.13	-.05	.06	-.01	-.05	.03	.00	.09
L_2 Linguistic Factors - Words		1.00	.21	.30	.29	.08	.62	-.04	.10	.00	.00	-.03	.26	-.06	-.01	.00
L_3 Linguistic Factors - Sentences			1.00	.24	.19	.08	.49	-.18	.06	-.12	-.18	-.09	.28	-.19	-.18	.00
L_4 Literal Comprehension 1.				1.00	.26	.14	.78	-.13	-.04	-.04	-.18	.07	.36	-.16	-.11	-.02
L_5 Interpretative Comprehension					1.00	-.01	.52	-.12	-.11	-.01	-.14	.14	.33	-.11	-.08	.03
L_6 Critical Comprehension						1.00	.41	-.18	.01	-.17	-.15	-.24	.05	-.42	-.23	.05
L_T Listening Total							1.00	-.20	.03	-.11	-.20	-.03	.42	-.29	-.19	.03
S_1 Voice Production								1.00	.14	.64	.33	.30	-.19	.97	.82	.06
S_2 Nonverbal									1.00	.19	.25	.09	-.15	.13	.37	.05
S_4 Linguistic Factors - Words										1.00	.46	.51	-.12	.64	.82	.12
S_5 Linguistic Factors - Sentences 2.											1.00	.30	-.31	.34	.66	.07
S_8 Organization of Details												1.00	.23	.34	.64	.05
SWT Speaking Written													1.00	-.19	-.16	-.08
R Ratings														1.00	.82	.04
S_T Speaking Total															1.00	.10
ATT Attitude																1.00

---

1. Contains two items from Domain III.

2. Contains one item from Domain VII.

a speech production is associated with vocal factors. As in Grades 3 and 6, there is a low negative relationship (-.19) between Speaking Total and Listening Total. We might conclude, again, that good speakers tend to be poor listeners.

Grade 12

Table 4-42 presents the correlation matrix for the Grade 12 Listening, Speaking, and Attitude data. The pattern is consistent with that for all other grades. The correlations among sub-scales in Listening are low, ranging from -.13 to .27. The sub-scale correlations with Listening Total (.20 to .69) are significant but generally low, suggesting considerable independence among the tests. Once again there is a low but consistent negative correlation between Listening and Speaking sub-scales. No significant correlations between Listening and Attitude appear.

With regard to Speaking the results are also similar to those found in other grades. Correlations among sub-scales are high, ranging from .29 to .82. Thus there is considerable dependence among these scales. This is expected because the ratings were all made by the same rater for each student and the halo effect would be operative. We note that there is a very low correlation between Speaking (Written) and all other scales. In fact, this sub-scale generally correlates more highly with the Listening Test than with the Speaking Tests. This suggests that the ability required to perform well on the written part of the test is unlike that required for speech production. The fact that the Ratings on Voice Production correlates .92 with Speaking Total suggests that we may be getting almost as much information from this sub-scale as we do from the total battery. Finally, we note that in terms of Attitude no significant correlations appear. Since this finding was consistent across all grades, we must seriously question the use of this instrument. More will be said of this later.

Table 4-42

Correlation Matrix for Listening, Speaking and Attitude by Domain Sub-scales and Total Test - Grade 12

		L_1	L_2	L_3	L_4	L_5	L_6	L_T	S_1	S_2	S_4	S_5	S_8	SWT	R	S_T	ATT
L_1	Voice Production 1.	1.00	.13	.07	-.04	.16	.03	.31	-.16	-.16	-.04	-.21	-.14	.16	-.17	-.13	-.12
L_2	Linguistic Factors - Words		1.00	.13	.20	.12	.17	.46	-.12	.00	-.06	-.04	-.02	.17	-.13	-.09	.02
L_3	Linguistic Factors - Sentences			1.00	.02	.16	-.13	.35	-.21	.01	-.14	-.04	-.08	.14	-.24	-.22	.02
L_4	Literal Comprehension				1.00	.27	.04	.69	.03	-.04	.11	-.05	-.03	.06	.00	-.01	-.07
L_5	Interpretative Comprehension					1.00	-.13	.66	-.21	.06	-.06	-.03	-.16	.22	-.25	-.20	-.05
L_6	Critical Comprehension						1.00	.20	-.01	-.09	-.03	-.05	.01	-.05	.02	.01	-.04
L_T	Listening Total							1.00	-.20	-.05	-.04	-.12	-.14	.23	-.24	-.19	-.08
S_1	Voice Production								1.00	.44	.63	.61	.55	-.10	.93	.92	.24
S_2	Nonverbal									1.00	.29	.82	.31	.12	.41	.44	.13
S_4	Linguistic Factors - Words										1.00	.52	.61	.17	.70	.75	.11
S_5	Linguistic Factors - Sentences											1.00	.52	-.03	.62	.62	.15
S_8	Organization of Details 2.												1.00	.20	.71	.76	.18
SWT	Speaking Written													1.00	-.15	.06	-.05
R	Ratings														1.00	.98	.29
S_T	Speaking Total															1.00	.28
ATT	Attitude																1.00

---

1. Contains two items from Domain III.

2. Contains one item from Domain VII.

FACTOR ANALYSIS

Since the domains, as noted above, were arrived at by an a-priori process, the Research Team felt that it would be useful to perform factor analyses on the Listening, Speaking and Attitude tests by grade. Test statistics, item analysis data and correlation data suggested that a male-female distinction in such an analyses would not be productive. The limited sample size, too, might lead to spurious results. Thus, a first-run factor structure based on the total population at each grade level was felt to be adequate and defensible. The results reported below are based on the use of the Varimax Rotation using an SPSS program.

For the purpose of analysis and factor interpretation a loading of .40 or over was considered arbitrarily to be significant. Further interpretation using lower loadings as cut-offs is, of course, possible. Similarly, other rotations might be attempted, but a reasonably good fit was obtained at most grade levels by this strategy.

Listening

Grade 3. Table 4-43 gives factor loadings for the Grade 3 Listening Test. Six factors, accounting for 76 per cent of the variance, were extracted. From Table 4-43 we can readily see that 10 items loaded at .40 or above on Factor I, 11 items on Factor II and so forth. For the purpose of interpretation we looked at an item mainly in terms of the factor on which it loaded most. For example, Item 13 loaded on Factors I, III and IV. In attempting to label

TABLE 4 - 43

Factor Structure for Grade 3 Listening Test

Item No.	Factor I	Factor II	Factor III	Factor IV	Factor V	Factor VI
1	-.05	.63*	.08	.04	.09	.13
2	.08	.64*	-.07	-.00	.07	-.24
3	.05	.76*	.01	-.09	.01	-.12
4	-.12	.76*	-.00	-.04	-.04	-.03
5	.03	.66*	-.02	-.03	.19	-.14
6	.03	.72*	.02	-.09	-.02	.25
7	-.10	.72*	-.13	-.08	-.16	.18
8	-.15	.53*	-.02	.09	.05	.16
9	.00	.59*	-.14	.04	.20	-.07
10	.09	-.14	.07	-.02	.20	.51*
11	.09	.86*	.07	.18	-.03	-.16
12	-.25	.22	-.29	-.25	.13	.00
13	.57*	-.05	.56*	.51*	.12	-.11
14	-.29	-.17	-.34	-.25	.10	.61*
15	.05	.08	.03	-.06	.88*	.26
16	-.12	-.21	-.14	-.16	-.83*	.00
17	.63*	-.07	.38	.06	.14	-.26
18	.54*	-.09	.69*	-.19	.15	-.09
19	.18	-.02	.21	.36	.05	-.04
20	.27	-.01	.68*	.66*	.05	-.04
21	.32	-.00	.79*	.42*	.00	.02
22	.63*	-.04	.64*	.29	.09	-.05
23	.70*	-.03	.28	.55*	.08	-.15
24	.25	.01	.07	.96*	.01	-.05
25	.27	-.01	.68*	.66*	.05	-.04
26	.25	.01	.07	.96*	.01	-.05
27	-.20	.03	.07	.20	-.00	.41*
28	.32	-.00	.79*	.42*	.00	.02
29	.51*	-.03	.30	.46*	.12	.01
30	.94*	-.01	.27	.07	.00	.03
31	.85*	-.00	-.05	.22	.03	.07
32	.94*	-.01	.27	.07	.00	.03
33	.79*	-.02	.33	.31	.05	-.09
34	.14	-.04	.91*	-.11	.06	-.00
35	.17	-.09	.81*	.06	.04	.00

the factors we initially placed it in Factor I. This is, of course, an arbitary procedure, but since reference factors were not available, we felt that it was an acceptable one. With these data before us, we now proceeded to attempt to label the factors and to match them with the domains that had been previously constructed. Table 4-44 presents the results of this analysis.

Inspection of Table 4-43 indicates that all of the items on the Grade 3 Listening Test can be placed in a six-factor space. Factor I, which contains 10 items, appears to be related to vocal characteristics of statements made in the test as well as to certain organizational characteristics. It seems likely that the speaker presents or stresses certain organizational patterns by the way he uses his voice. Factor II has been tentatively labelled as one involving Interpretative Communication and Content Relationships. It includes abilities to infer feelings, to find the main idea, supporting details and organizational patterns, and to derive alternative hypotheses. Factor III has been labelled as Relations and Going Beyond the Information given. For the most part it is related to drawing conclusions and using structural patterns of a message to gain information. In each case there are, of course, items which do not seem to fit the pattern, but in general the first three factors appear to be fairly meaningful. Factors IV, V, and VI are less clear. Since only two items load on each it is difficult to interpret them.

The next step in the procedure was to compare the factor structure to the original or a-priori domains. We note from Table 44 that Factor I contains four items from Domain I, one item from Domain 2, two items from Domain 4, and one each from Domains 5 and 6. At first glance this appears to be a rather tangled mixture. If the factor structure has validity, however, it can be seen that the items tend to measure abilities related to deriving meaning from characteristics of vocal patterns. That is, meaning is

TABLE 4-44

Factor Labels and Comparison with A-priori Domains

Grade 3 Listening

FACTOR	ITEM NO.	DOMAIN	OBJ. NO.	OBJECTIVE DESCRIPTION
1				
Interpretive Communication/ Voice Production Factors	13	5	5C2H	speaker's purpose
	17	4	5C1E	beginning, middle ending sequence
	19	2	4B1	meanings of denotative words
	23	1	2D1	vocal quality and tone of voice
	29	6	5C30	draw conclusion
	30	1	2B1	intonation
	31	1	2B1	intonation
	32	1	2D1	vocal quality and tone of voice
	33	1	2D1	vocal quality and tone of voice
2				
Interpretive Communication Content Related Factors	1	5	5C2C	infer the thought, feelings, emotions
	2	5	5C2C	infer the thought, feelings, emotions
	3	5	5C2F	alternative hypotheses
	4	4	5C1C	supporting details
	5	4	5C1A	main idea
	6	4	5C1E	beginning, middle ending sequence
	7	4	5C1E	beginning, middle ending sequence

TABLE 4-44 continued

FACTOR	ITEM NO.	DOMAIN	OBJ. NO.	OBJECTIVE DESCRIPTION
2				
cont'd.	8	5	5C2F	alternative hypotheses
	9	4	5C1H	organizing pattern
	11	1	2E1	pace
3				
Ideational Communication/Beyond Given Information	18	2	4B7	variety of types of words
	20	6	5C30	draw conclusion
	21	1	2E1	pace
	22	2	4B7	variety of types of words
	25	5	5C2F	alternative hypotheses
	28	4	5C1H	organizing pattern
	29	6	5C30	draw conclusion
	34	1	2D1	vocal quality and tone of voice
	35	1	2D1	vocal quality and tone of voice
4				
	24	6	5C30	draw conclusion
	26	5	5C2H	speaker's purpose
5				
	15	4	5C1C	supporting details
	16	2	4B1	meaning of denotative words

TABLE 4-44 continued

FACTOR	ITEM NO.	DOMAIN	OBJ. NO.	OBJECTIVE DESCRIPTION
6				
	10	6	5C3D	fact or fiction
	12	4	5C1A	main idea
	14	2	4B1	meanings of denotative words
	27	6	5C3D	fact or fiction

derived in such items as 9, 13, 17, 19, and 27 from the way in which the speaker presents the message. For example, it may be that students recognize beginning, middle, and ending from the stress and intonation placed on the words by the speaker.

Factor II derives largely from Domains 4 and 5 and is largely a factor related to literal and interpretive comprehension. There is obviously some overlap with Factor I, but the match in terms of original domains is quite good.

Factor III is a mixture once again and is difficult to relate to the original domain structure. It is likely here that this factor is based mainly on residuals from the first two and that at this point no meaningful comparison is possible. The same is true for Factors IV, V, and VI where the number of items involved is, in any case, too small to make comparisons feasible.

Grades 6-12. A similar factor analysis was applied to the Grade 6, 9, and 12 data in Listening. Although a number of rotations were attempted, no meaningful factor structure could be obtained. In all cases at least 20 factors were required and even then only about half of the items could be accounted for. In addition, only two to three items loaded substantially on any one factor. As a result no tables providing these data are presented.

The problem is, of course, to account for these findings. Several possibilities exist. First of all, it may be that a criterion-referenced test, measuring as it does a heterogeneous sample of abilities, may just not lend itself to this kind of analysis. In other words, we are, in fact, measuring a set of distinct abilities which are not contained in a limited factor space. The question then arises as to why the Grade 3 results should provide a factor structure when others do not. It should be noted, however, that the Grade 3 test differs in nature from the others. (See p. 3-30)

It is also possible that abilities at lower grade levels are more homogeneous by nature and that a number of skills do not operate significantly until a certain level of maturity is reached.

A second possibility resides in the statistical procedures themselves. The data in the test are dichotomous in nature. While most programs will use these data to produce Pearson Product Moment Correlations, which form the basis of the factor analysis, it may be that tetrachoric correlations would be more suitable. It is possible that with larger sample sizes and tetrachoric correlation coefficients, a logical structure might be found.

In sum, this problem requires further study with a new and larger sample. Although it can be argued, we feel that the difficulty in producing a reasonable factor space should not be seen as a negative finding for a criterion-referenced test. In fact, if the items are valid and measure important objectives, it can be argued that these findings are positive.

Speaking

Grade 3. Table 4-45 gives factor loadings for the Grade 3 Speaking Tests. Six factors accounted for 90 percent of the variance. Again a loading of .40 or higher was used as the basis for analysis. It can be seen that 10 items loaded significantly on Factor I, II on Factor III, etc. For the purposes of initial analysis and the labelling of factors we once again looked at an item as being mainly in a factor where its loading was the highest.

Table 4-46 lists the items that fall within each factor. Factor I appears to relate to the external manifestations of a speech production and the way in which sentences are put together to form

TABLE 4 - 45

Factor Structure for Grade 3 Speaking Test

Item No.	Factor I	Factor II	Factor III	Factor IV	Factor V	Factor VI
1	.15	.71*	.05	-.07	-.00	.14
2	.29	.64*	.18	[illegible]	.04	-.00
3	-.03	.47*	-.37	.08	-.22	.20
4	.05	.70*	.03	-.04	-.02	.38
5	.17	.71*	.08	.03	-.05	-.13
6	.11	.55*	.29	-.13	.36	.05
7	.05	.45*	.18	-.11	-.00	.26
8	.31	.54*	.16	.07	.05	.04
9	.50*	.61*	.21	-.11	.04	-.06
10	.14	-.74	.24	.11	.11	-.15
11	.01	.21	.40*	.07	-.10	.22
12	.13	.12	.87*	-.08	.10	.09
13	.28	-.07	.70*	.03	.06	-.03
14	.23	-.02	.76*	-.01	-.01	-.02
15	.19	.11	.55*	-.00	.06	.16
16	.02	.13	.20	-.08	.01	.05
17	.14	.18	.68*	-.08	.03	.07
18	-.02	.28	.18	.02	.01	.04
19	-.13	.41*	.00	.02	.05	.24
20	.69*	.02	.38	.03	.14	.05
21	.70*	.22	.31	-.02	-.02	-.08
22	.71*	.18	-.02	-.07	.02	.09
23	.55*	.12	.14	.00	-.01	.51*
24	.67*	.06	.08	-.04	.09	.43*
25	.64*	.21	.00	.07	.10	.09
26	.43*	.26	.09	-.08	.04	.66*
27	.80*	.09	.29	.01	.09	.04
28	.73*	-.20	.23	.05	.13	-.11
29	-.06	.01	-.05	-.02	-.61	-.06
30	-.04	.12	-.01	.74*	.03	.14
31	.02	.10	-.02	.30	.13	.01
32	-.04	-.10	-.08	.56*	.08	-.17
33	-.03	-.09	.08	-.14	.40*	.01
34	.03	-.21	-.01	.40*	-.01	.04
35	-.14	.00	.01	.09	-.60	.14
36	.06	-.00	.02	.76*	.07	-.07
37	-.04	.04	.19	.11	.55*	.09
38	-.11	.14	-.02	-.11	-.40	.00

TABLE 4-46

Factor Labels and Comparison with A-priori Domains

Grade 3 Speaking

FACTOR 1	ITEM NO.	DOMAIN	OBJ. NO.	OBJECTIVE DESCRIPTION
Oral Speech Production - External Measures	20	4	3B2	variety of types of words
	21	5	3C8	complete sentences
	22	5	3C7	syntax
	23	5	3C7	syntax
	24	5	3C8	complete sentences
	25	6	4A1	associations in the form of ideas
	26	8	4C6	inferences
	27	1	1E2	no hesitations
	28	1	1E4	length of time
2				
Oral Speech Analysis	1	8		
	2	4	3B6	connotative meaning
	3	4	3B5	denotative meaning
	4	4	3B5	denotative meaning
	5	8	4C6	draw inferences from details

TABLE 4-46 continued

FACTOR 2	ITEM NO.	DOMAIN	OBJ. NO.	OBJECTIVE DESCRIPTION
cont'd.	6	8	4C1	main idea
	7	8	4C5	sequence
	8	8	4C5	sequence
	9	8	4C1	main idea
	10	1	1D1	where to pause
	19	3	3A1	where to pause
3 manifest speech production	11	1	1C2	vocal stress
	12	1	1E3	pace
	13	2	2B1	eye contact
	14	2	2B2	body movement
	15	4	3B2	pronunciation
	16	8	4C6	inferences
	17	8	4C2	exclude details not supporting main idea
	18	8	4C4	develop beginning, middle & ending
	37	3	3A2	pronunciation

TABLE 4-46 continued

FACTOR	ITEM NO.	DOMAIN	OBJ. NO.	OBJECTIVE DESCRIPTION
4				
vocal characteristics	31	1	1C1	natural voice
	32	1	1C2	vocal stress
	33	1	1D2	does pause
	34	1	1E3	pace
	20	1	1A1	project voice
5				
General Presentation Factors	35	2	2B2	body movements
	38	1	1E2	no hesitations
	36	3	3A1	articulates speech sounds

a communication. Factor II relates to oral speech analysis or organization which is internal to the speaker himself. Basically, it involves the way words or phrases are chosen and organized in order to produce meaning. Factor III involves primarily those aspects of a speech which are manifest to others. This includes pronunciation and non-verbal characteristics of the speaker. It is not immediately apparent why Items 16, 17, and 18 which have to do with organization of details, should fall into this category unless speakers in some way provide vocal cues to indicate such organization. Factor IV relates primarily to vocal characteristics such as stress, pause and projection. Factor 5, which contains only four items is difficult to label but appears to contain items relating to general presentation factors. This is also true of Factor IV on which no items show a primary loading.

Inspection of Table 4-46 also indicates the match between the factor structure and the original domains. Factor I, labelled as Manifest Speech Production, includes mainly items from Domain 4, which has to do with Linguistic Factors - Sentences, together with some voice production items and fluency. Factor II, labelled as Oral Speech Analysis, consists primarily of items from Domains 4 and 8, which relate to Linguistic Factors - Words and Organization of Details. Thus choice of words and sequencing of details appear to be related. Factor IV - Vocal Characteristics - consists of items only from Domain 1, Vocal Production, and is therefore a perfect match. Since Factor V contains only three items, it cannot be labelled. The same is true for Factor VI.

Grade 6. Table 4-47 gives factor loadings for the Grade 6 Speaking Tests. Six factors accounted for 93 percent of the variance. It should be noted that Items 1-15, the items on the written test, load heavily on Factor IV, but they also loaded substantially, and in a negative direction, on Factor I. It is difficult to explain this finding on the basis of present information. At this point it seems best to look at these items as a part of

TABLE 4-47

Factor Structure for Grade 6 Speaking

ITEM	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6
1	-.42*	.12	-.14	.74*	.24	-.04
2	-.45*	.19	-.11	.66*	.07	-.03
3	-.51*	.08	-.13	.72*	.13	.00
4	-.46*	.09	-.16	.76*	.15	.02
5	-.39	.03	-.12	.78*	.21	-.02
6	-.43*	.15	-.14	.75*	.21	-.01
7	-.46*	.10	-.12	.77*	.01	.03
8	-.42*	.12	-.20	.68*	.05	-.01
9	-.48*	.14	-.06	.75*	.04	.03
10	-.41*	.19	-.03	.76*	.02	-.05
11	.45*	-.19	.11	-.53*	-.12	.06
12	-.35	.09	-.13	.79*	.19	.01
13	-.43*	.03	-.14	.47*	-.04	-.03
14	.10	-.10	.11	-.57*	-.19	.00
15	.25	-.01	.04	-.55*	-.10	.01
16	-.08	.03	.11	.14	.76*	.05
17	-.00	.12	-.09	.37	.69*	.06
18	.31	-.14	.25	-.53*	-.12	.27
19	.28	-.00	.33	-.49*	-.00	.36
20	.29	.06	.43*	-.36	.13	.32
21	.19	.04	.42*	.09	.31	.33
22	.02	.18	-.03	.15	.30	.44*
23	.18	.19	-.07	-.13	.05	.53*
24	-.02	.46*	.02	.34	.35	.31
25	-.30	.31	-.09	.65*	.18	-.00
26	.23	.10	.76*	-.36	-.03	-.23
27	.27	.10	.65*	-.42*	-.02	-.06
28	.07	.84*	-.02	.08	-.12	.11
29	-.06	.75*	-.08	.15	.19	.02
30	-.08	.76*	-.08	.27	.08	.04
31	.13	.67*	.14	-.10	.04	-.00
32	-.16	.44*	.27	-.03	.04	.03
33	-.06	.57*	.04	.34	-.00	.10
34	.30	.47*	.34	-.14	.00	-.05
35	.38	.21	.33	-.50*	-.17	.03

TABLE 4-47 continued

ITEM	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6
36	.66*	.10	.08	-.38'	.03	.18
37	.69*	.16	.01	-.51*	-.00	.03
38	.70*	.01	.07	-.43*	-.08	.07
39	.68*	-.05	.19	-.40*	.01	.13
40	.80*	-.06	.19	-.35	.02	.00
41	.78*	-.08	.14	-.32	-.08	-.01
42	.65*	-.05	.13	-.47*	-.11	.11
43	.76*	.03	.05	-.40*	.01	.11
44	.77*	.02	.13	-.34	-.00	.13
45	.83*	.00	.08	-.36	.01	.09

Factor IV where they have the heaviest loading.

Factor I has been labelled Oral Presentation Factors - Manifest Characteristics and consists mainly of voice production, pronunciation, articulation, fluency and non-verbal items. Factor II has been labelled a Fluency Factor. Items 28-31, which may be seen as related to syntax, nevertheless can be construed as being related to fluency as can sequencing, Items 33. Factor III relates purely to items which measure variety. It probably measures that characteristic of a speech which makes it interesting. Factor IV we have labelled as General Ability and Observed Speech Production - that is, it involves elements of speaking which derive from the speaker's background of information and his knowledge about how a communication should be produced. It is in this factor that knowledge of how to produce a speech, as measured by the written test, comes into play. Factors V and VI have not been labelled.

Table 48 indicates the relationship between the factor structure and the a-priori domains. Factor I derives from items from Domains 1, 2, and 3. Since Domains 1, 2, and 3 were related to voice production, non-verbal factors and sounds, this appears to be a close match. Factor II, Fluency, relates mainly to Domain 5, Sentences. Factor III, Variety, relates to Domains 4 and 5, Linguistic Factors - Words and Linguistic Factors, Sentences. Factor IV, General Ability and Observed Speech Production represents mainly Domain 1, Voice Production and Domain 4, Linguistic Factors, Words. Factors V and VI relate to Domains 1 and 8 respectively, but since these factors were not labelled, a comparison can be suggestive only.

Grade 9. Table 4-49 gives factor loadings for the Grade 9 Speaking Tests. Six factors accounted for 96 percent of the variance. Items were again assigned to factors on the basis of their heaviest loading.

TABLE 4-48

Factor Labels and Comparison with A-priori Domains

Grade 6 Speaking

FACTOR	ITEM NO.	DOMAIN	OBJ. NO.	OBJECTIVE DESCRIPTION
1				
Oral Speech	36	1		project voice
	37	1	1A1	projects voice
	38	1	1C1	natural voice
	39	1	1C2	vocal stress
	40	1	1D2	does pause
	41	1	1E3	pace
	42	2	2B2	body movements
	43	3	3A1	articulates speech sounds
	44	3	3A2	pronunciation
	45	1	1E2	no hesitations
2				
Linguistic Factors - Fluency	24	8	4C4	develop beginning, middle & ending
	28	5	3C7	syntax
	29	5	3C7	syntax
	30	5	3C7	syntax
	31	5	3C8	complete sentences

TABLE 4-48 continued

FACTOR 2	ITEM NO.	DOMAIN	OBJ. NO.	OBJECTIVE DESCRIPTION
cont'd.	32	6	4A1	associations in form of ideas
	33	8	4C5	sequence
	34	1	1E2	no hesitations
3				
Linguistic Factors - Variety	20	4	3B2	variety of types of words
	21	5	3C2	variety of sentence levels
	26	4	3B2	variety of types of words
	27	5	3C4	variety of transitional devices
4				
General Ability and Observed Speech Production	1	7	4B5	express meaning, mood or tone
	2	4	3B5	denotative meaning
	3	4	3B6	connotative meaning
	4	4	3B6	connotative meaning
	5	1	1A3	point stressed
	6	8	4C6	inferences

TABLE 4-48 continued

FACTOR 4	ITEM NO.	DOMAIN	OBJ. NO.	OBJECTIVE DESCRIPTION
cont'd	7	1	1D1	where to pause
	8	1	1E1	ways rate effects, meaning, etc.
	9	1	1A3	point stressed
	10	1	1A4	indicate ways volume, pitch, rate & quality affect meaning
	11	1	1E1	rate - meaning, mood, etc.
	12	4	3B5	denotative meaning
	13	8	4C6	inferences
	14	1	1D1	does pause
	15	7	4B5	express meaning, mood, tone
	18	2	2B1	eye contact
	19	2	2B2	body movements
	25	3	3A1	articulate speech sounds
	35	1	1E4	length of time

TABLE 4-48 continued

FACTOR	ITEM NO.	DOMAIN	OBJ. NO.	OBJECTIVE DESCRIPTION
5				
Vocal Quality	16	1	1C2	vocal stress
Rate, Time Order	17	1	1E3	pace
6				
Time Order	22	8	4C5	sequence
Irrelevant Details	23	8	4C2	exclude details not supporting main idea

TABLE 4-49

Factor Structure for Grade 9 Speaking

ITEM	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6
1	-.36	-.61*	-.10	.11	-.29	-.06
2	-.39	-.63*	-.10	.09	-.33	-.01
3	-.27	-.77*	-.10	.04	-.25	-.06
4	-.36	-.73*	-.09	-.02	-.29	.04
5	-.34	-.73*	.00	-.03	-.35	.06
6	-.34	-.68*	-.07	.01	-.30	.06
7	-.32	-.53*	-.10	.03	-.46*	.05
8	-.28	-.80*	-.07	.06	-.20	.04
9	-.24	-.77*	-.03	.06	-.23	.10
10	.38	.47*	.05	-.08	.56*	-.08
11	-.32	-.83*	-.07	-.02	-.20	-.02
12	-.29	-.31	-.10	.06	-.31	.05
13	.44*	.48*	.07	-.09	.56*	-.12
14	.45*	.49*	.03	-.09	.60*	-.14
15	.41*	.46*	.09	-.07	.55*	-.07
16	.33	.26	.03	-.12	.66*	-.12
17	.47*	.36	.17	-.10	.55*	-.16
18	.33	.25	-.03	-.06	.70*	-.06
19	.35	.27	-.04	-.14	.37	-.14
20	.40*	.32	.09	-.09	.44*	.03
21	.32	.33	.03	-.12	.73*	-.09
22	.41*	.41*	.09	-.13	.67*	-.11
23	.31	.31	.06	-.12	.46*	-.06
24	.74*	.33	.11	-.18	.30	-.09
25	.62*	.38	.16	.00	.26	-.02
26	.70*	.30	.18	.02	.24	-.06
27	.63*	.42*	.03	-.04	.30	-.15
28	.64*	.44*	.03	-.06	.30	-.08
29	.69*	.38	.02	-.05	.33	-.01
30	.70*	.25	-.03	-.00	.39	.01
31	.66*	.26	.04	-.05	.34	.06
32	.70*	.37	.21	-.03	.29	.01
33	.70*	.31	.24	-.02	.25	-.04
34	.72*	.29	.18	-.10	.30	-.06
35	.69*	.36	.09	-.03	.29	.02

TABLE 4-49 continued

ITEM	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6
36	.32	.15	.33	.50*	.09	.02
37	-.26	-.11	.00	.54*	-.27	.10
38	-.21	-.20	.03	.69*	-.15	.14
39	.05	.05	.19	.15	-.17	.70*
40	-.32	-.21	-.09	.34	-.29	.50*
41	.27	.09	.15	.51*	.09	.13
42	.41*	.18	.38	.26	.30	-.02
43	.55*	.38	.13	.15	.44*	-.02
44	.52*	.41	.13	.17	.45*	-.03
45	.40*	.36	.11	.23	.38	-.05
46	-.45*	-.29	.24	.49*	-.33	-.09
47	-.04	.01	.75*	.03	.01	.16
48	.30	.25	.81*	.00	.14	-.02
49	-.27	-.20	.27	.50*	-.33	-.03
50	.31	.13	.74*	.13	.10	.01
51	-.05	-.09	.68*	.27	-.11	-.03
52	.08	.18	.24	.32	-.03	-.26
53	-.53	-.39	.08	.20	-.55*	.03
54	-.50	-.42	.07	.14	-.57*	.02

3 0

Table 4-50 indicates the items and the factors on which they had their primary loading. Factor I appears to describe speaking style or manner of presentation. It involves those characteristics which a rater might look for as the immediate indicators of a good speech. Factor II has been labelled General Analytical Ability. It appears to involve the extent to which the speaker is able to react to or take into account observed audience response and to use this information to make decisions about how a message should be structured. This factor does not seem to be evident in Grades 3 and 6 suggesting that at lower levels the student is not able to perform this activity specifically. It should be noted, however, that these items were all on the written test and that the relationship between knowledge of how a speech should be presented and the ability to present it effectively can only be determined from the correlation among these factors. Factor III represents Vocal Quality and Pace. Factor IV does not appear too different from Factor III. It is made up of some linguistic items, fluency items and items related to the ability to project and to pause correctly. The loading of these items seems to suggest that raters are able to discriminate between simple vocal quality and those vocal qualities which give meaning and emphasis to a spoken message. Factor V relates to preorganization ability, diction and looking ahead. These items measure the ability of the student to signal to an audience what his major message is. To the extent that he knows his direction, he is then able to present his message smoothly and establish order among the details presented. Item 53, Diction, also falls into this factor. This may suggest that there is a relationship between thought and deed in that the student who knows what he wants to say can do so clearly. Factor VI, though it includes only two items, is interesting because sequencing and using transitional devices seem to be a distinct ability.

TABLE 4-50

Factor Labels and Comparison with A-priori Domains

Grade 9 Speaking

FACTOR 1	ITEM NO.	DOMAIN	OBJ. NO.	OBJECTIVE DESCRIPTION
Speaking Style	24	1	1C1	natural voice
	25	1	1E3	pace
	26	3	3A1	articulate speech sounds
	27	2	2A2	avoid distracting movements
	28	2	2B2	body movements
	29	5	3C2	sentence levels
	30	5	3C8	complete sentences
	31	8	4C2	exclude irrelevant details
	32	8	4C4	develop introduction body & conclusion
	33	8	4C4	develop introduction body & conclusion
	34	4	3B3	words & expressions appropriate to language level
	35	7	4B3	state purpose
	42	1	1E4	length of time
	43	5	3C5	stylistic devices
	44	5	3C5	stylistic devices
	45	5	3C5	stylistic devices

TABLE 4-50 continued

FACTOR 2	ITEM NO.	DOMAIN	OBJ. NO.	OBJECTIVE . DESCRIPTION
General Analytical Ability	1	8	4C6	infer meaning
	2	1	1A3	point stressed
	3	1	1E1	indicate how rate, affects meaning
	4	1	1D1	where to pause
	5	1	1D1	where to pause
	6	1	1B1	intonation
	7	1	1A3	point stressed
	8	2	2B3	eye contact
	9	1	1E1	indicate how rate, affects meaning
	10	8	AC6	infer meaning
	11	5	3C6	stylistic devices
	12	4	3B5	denotative meaning
	13	4	3B6	connotative meaning
	14	4	3B6	connotative meaning
	15	8	4B6	infer meaning
	16	8	4C6	draw inferences

TABLE 4-50 continued

FACTOR	ITEM NO.	DOMAIN	OBJ. NO.	OBJECTIVE DESCRIPTION
3				
Vocal Quality and Pace	47	1	1C1	natural voice
	48	1	1C2	vocal stress
	50	1	1E5	pace
	51	1	1E3	pace
4				
Linguistic Factors - Fluency, Projection, Pause	36	4	3B2	variety of types of words
	37	5	3C7	syntax
	38	5	3C7	syntax
	41	1	1E2	no hesitations
	46	1	1A1	project voice
	49	1	1D2	does pause
5				
Pre-Organization Diction, Looking Ahead	17	8	4C1	main idea
	18	8	4C4	develop introduction body & conclusion
	19	8	4C4	develop introduction body & conclusion
	20	8	4C9	select and develop topic sentence
	21	8	4C9	select and develop topic sentence

3.4

TABLE 4-50 continued

FACTOR 5	ITEM NO.	DOMAIN	OBJ. NO.	OBJECTIVE . DESCRIPTION
cont'd.	22	8	4C5	sequence
	23	8	4C5	sequence
	53	3	3A1	diction
	54	1	1E2	no hesitations
6				
Variety of Connectors and does not Return	39	5	3C4	transitional devices
	40	8	4C5	sequence

Table 4-50 also indicates the extent to which there is a relationship among the factors and the original domains. Factor I contains items from almost all of the domains, being most heavily represented by Domains 5, 1, and 8. If "Speaking Style" is an adequate label for Factor I, this suggests that such style is largely determined by the kinds of sentences used, certain vocal qualities, and the way in which details are organized. Factor II contains seven items from Domain I, three from Domain 4, three from Domain 8, and one each from Domains 2 and 5. No clear comparison emerges, but as already pointed out, this factor appears to load on items that require analysis of how a speech should be structured and they do appear to hang together. Factor III loads only on items from Domain I and is clearly related to voice production. Factor IV is made up of items from Domains 1, 4, and 5. It appears to relate mainly to what we originally structured as Linguistic Factors - Words and Sentences. Items contained in Factor V came largely from Domain 8, Organization of Details. Items 53 and 54 are the exception, but it seem reasonable to suggest that a clear organization of details might lead to clear diction and fluent presentation as measured by these items. Factor VI, because it contained only two items, cannot appropriately be compared in this analysis.

Grade 12. Table 4-51 presents the factor loading for the Grade 12 Speaking Test. In this case six factors accounted for 76 percent of the variance in the Varimax Rotation. We note that this figure is substantially lower than that for the other grades. This is in large part accounted for by the fact that the written part of the test did not load on any specific factor; in fact, many items did not load on primary factors at all. In this case the written test items behaved much as did the items on the Listening Test. That is, they scattered to many smaller factors. Two conclusions may be drawn from this result. First of all, the items in Speaking (Written) are not homogeneous with respect to ability. They appear

TABLE 4-51

Factor Structure for Grade 12 Speaking

ITEM	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6
1	.04	-.04	.00	-.25	.38	.00
2	-.09	-.00	.00	-.08	.21	-.06
3	.05	-.05	.10	.04	.04	.18
4	.13	-.23	.17	.04	-.03	-.14
5	-.01	-.00	.02	-.01	.00	-.02
6	.04	.01	.03	.05	.18	.05
7	-.03	.08	-.11	-.25	.23	.02
8	-.06	-.14	.01	.03	.45*	.15
9	.19	-.04	.02	-.07	.22	-.01
10	-.01	-.05	.07	-.00	.08	.01
11	.08	.01	.03	-.04	-.04	-.05
12	.01	-.02	-.07	-.09	.42*	-.01
13	.00	-.07	.03	.03	.51*	-.08
14	-.11	.02	.01	.05	.07	-.06
15	-.05	-.05	.00	-.05	-.05	-.00
16	-.13	-.06	.06	.01	-.00	-.07
17	-.20	-.14	.03	.02	.10	.01
18	.01	-.12	-.09	.02	.05	.07
19	.05	-.02	-.00	.05	.10	.13
20	.23	.63*	.02	.19	-.23	.02
21	.06	.76*	.10	.22	-.07	.09
22	.12	.74*	.09	.15	-.11	-.01
23	.16	.45*	.13	.04	-.09	.58*
24	.18	.44*	.00	-.01	.05	.72*
25	.28	.66*	.25	.02	.02	.22
26	.24	.70*	.11	.09	.08	.10
27	.11	.74*	-.09	.17	.05	.10
28	.21	.65*	.07	.08	-.12	.08
29	.14	.59*	.17	.02	-.14	.01
30	.18	.79*	.15	-.02	-.13	.05
31	.10	.62*	-.04	.18	-.12	.23
32	.24	.45*	.16	.50*	-.07	-.07
33	.09	.36	.50*	.40*	.12	.03
34	.13	.30	.43*	.40*	.07	-.05
35	.22	.42*	.37	.44*	-.01	.00

TABLE 4-51 continued

ITEM	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6
36	.14	.28	.13	.74*	-.12	.01
37	.12	.32	.27	.65*	-.01	.05
38	.28	.42*	.24	.19	.26	-.07
39	.12	.29	.45*	.10	-.44*	.11
40	.11	.06	.78*	.07	-.05	.02
41	.05	.06	.81*	.13	.01	.03
42	.04	.15	.79*	.09	.01	-.01
43	.44*	.36	.15	.20	.00	.03
44	.51*	.37	.04	.06	-.08	.02
45	.73	.25	.10	.17	-.01	.02
46	.76*	.20	.14	.11	-.02	.07
47	.83*	.21	.03	.04	-.07	.11
48	.78*	.29	.09	.02	.06	.08
49	.42*	.08	-.13	.16	.05	.16
50	.43*	.46*	.07	.01	.09	-.15
51	.49*	.51*	.05	.03	.09	-.25

to require rather different abilities in each case, and this ability is probably related to each student's background. Secondly, the dispersion we observe suggests that as students get older they depend more on specific knowledge to answer questions of this kind -- no general ability factor is operating.

We note also that the remaining items group themselves rather firmly on the basis of the Stimulus Sub-scale. For example, items 43 to 51 were based on ratings given during the oral reading of a poem. This grouping is probably largely due to halo effect in rating. Further study would indicate whether the number of items might be reduced, but such a reduction would require that we also ignore certain objectives.

Table 4-52 outlines the factor structure of the Grade 12 Speaking Test and compares this to the original domains. Factor I contains items almost exclusively from Domain 1 - Voice Production. The vocal aspects of a speech production thus form the major basis for rating quality. Factor II contains items from almost all of the domains. It is difficult to label this factor, but it appears that once again we have a group of items which describe what a rater perceives as the outward mainfestations of a speech or reading. Thus fluency, diction, smoothness of presentation, language level and general verbal ability group themselves together to reflect a speaker's agility in presentation. It is, of course, also possible that what we have here is the rater's stereotype of what he/she expects in a speech production -- that is, the rater may tend to group certain qualities and rate them similarly. Factor III matches with a section of Domain 5. This has been labelled Sentence Meaning and Structure. Factor IV appears to define what might be called Coherence. It involves the ability to present material in concise form and to present ideas in a logical way. Factor V contains items referring to variety and transition, while Factor VI includes two items on the non-verbal aspects of speaking behaviour.

TABLE 4-52

Factor Labels and Comparisons with A-priori Domains

Grade 12 Speaking

FACTOR	ITEM NO.	DOMAIN	OBJ. NO.	OBJECTIVE DESCRIPTION
1				
Voice Production	43	1	1A1	projection
	44	1	1C1	natural voice
	45	1	1C2	varies vocal quality
	46	1	1D2	pause
	47	1	1E5	pace
	48	1	1E3	pace
	49	1	2B2	nonverbal
2				
Manifest Speech Production	20	1	1C1	natural voice
	21	1	1E3	pace
	22	3	3A1	projection
	25	5	3C2	variety of sentence levels
	26	5	3C8	syntax
	27	8	4C2	excludes irrelevant details
	28	8	4C4	introduction body & conclusion
	29	8	4C4	introduction body & conclusion
	30	4	3B3	appropriate language level

TABLE 4-52 continued

FACTOR	ITEM NO.	DOMAIN	OBJ. NO.	OBJECTIVE DESCRIPTION
2				
Manifest Speech Production continued	31	7	4B3	appropriate language level
	32	4	3B2	variety of words
	35	5	3C4	transitional devices
	38	1	1E2	fluency
	50	3	3A1	projection
	51	1	1E2	fluency
3				
Sentence Meaning and Structure	33	5	3C7	stylistic devices - meaning mood
	34	5	3C7	stylistic devices - meaning mood
	40	5	3C5	variety of stylistic devices
	41	5	3C5	variety of stylistic devices
	42	5	4C5	sequence-details
4				
Coherence	32	4	3B2	variety of word types
	35	5	3C4	transitional devices
	36	8	4C3	sentence types
	37	8	4C5	details-organization

TABLE 4-52 continued

FACTOR	ITEM NO.	DOMAIN	OBJ. NO.	OBJECTIVE DESCRIPTION
5				
Variety and Transition	8	4	3B6	denotative meanings
	12	8	4C1	sentences - express meaning mood etc.
	13	8	4C4	sequence details
	39	1	1E4	speak for prescribed length of time
6				
Nonverbal	23	2	2A2	nonverbal
	24	2	2B2	nonverbal
Items not appearing in first 6 Factors	1	1	1D1	pause
	2	1	1E1	rate
	3	4	3B5	voice production - meaning mood
	4	1	1D1	pause
	5	7	4B5	inferencing
	6	7	4B5	inferencing
	7	4	3B6	denotative meanings
	9	4	3B6	denotative meanings
	10	2	2B3	nonverbals - audience contact
	11	1	1A4	voice production - meaning mood

TABLE 4-52 continued

FACTOR	ITEM NO.	DOMAIN	OBJ. NO.	OBJECTIVE DESCRIPTION
continued	14	8	4C4	introduction body & conclusion
	15	8	4C9	topic selection and development - sentence
	16	8	4C9	topic selection and development - sentence
	17	8	4C9	topic selection and development - sentence
	18	8	4C5	organizational pattern
	19	8	4C5	organizational pattern

Unlike the results from the factor analysis at Grades 3, 6, and 9, we note that 16 items do not appear in the factor structure. All of these items appeared in the written test. As noted above, we feel that this is probably the case because at this level the written questions require a rather heterogeneous knowledge based on experience and opinion rather than on some general educational factor to which all students have been exposed in the course of their school work.

ATTITUDE

As pointed out earlier in this report, one of the major concerns about the process of speaking is how to deal with the anxiety which arises in formal communication situations. It was of interest, then, to give a test which would measure communication apprehension. For this purpose the Personal Report of Communication (PRCF), a questionnaire consisting of 14 items requiring responses on a 5-point scale, was administered to Grades 3 and 6. The Personal Report of Communication Apprehension (PRCA), consisting of 25 items requiring ratings on a 5-point scale, was administered to Grades 9 and 12.

According to the author of the tests, scores between 28 and 47 and between 62 and 88 are considered to in the normal range for the two tests respectively. Students who fall above or below these points in the scale may have difficulties in school. Those whose self-reports place them above the norm are considered to be fearful about communicating, while those who fall below the norm are likely to be so verbally active as to create discipline problems for teachers and social problems for themselves.

Table 4-53 presents the means and standard deviations by grades on the PRCF and PRCA. Based on the norms as given by the test constructor, we would conclude that for this sample, at the

TABLE 4-53

Means and Standard Deviations for PRCF and PRCA

Grade	Test	Mean	S.D.	Normal Range
3	PRCF	34.59	8.57	28 - 47
6	PRCF	34.78	9.01	28 - 47
9	PRCA	73.77	13.64	62 - 88
12	PRCA	73.60	14.88	62 - 88

Grade 3 and 6 level, very few--roughly one percent--of the students suffer from communication apprehension. On the other hand, some four percent are highly verbal on the basis of self report. The results on the PRCA are almost identical. We note also that the means for the two grades on each test are almost identical, suggesting that if the test is valid, there is no change in level of anxiety from one grade level to the next. From the available norms we can only conclude that the test probably does not allow us to isolate students with this kind of communication problem. It would have been interesting to compare the test scores with ratings made by teachers or others who knew the students well. We were unable to get such information in this study. Furthermore, since the attitude questionnaire did not correlate with the Speaking and Listening Tests at any grade level (See Tables 4-39 to 4-42), its validity as a useful measure must remain in question.

Chapter 5

CONCLUSIONS AND RECOMMENDATIONS

On the basis of the data analysis the Research Team concluded that, with minor revisions, the field-tested instruments could now be used as Final Tests for large-group testing in Alberta schools. It should be noted, however, that these Final Tests, especially the Speaking Tests, are relatively novel, as was the developmental approach generally taken by the Research Team. For this reason, we make a number of recommendations about further test development. The next section summarizes our conclusions and provides suggestions for the use of the tests as well as for further test development.

Future Test Administration

Conclusion 1: Item analysis and other data indicate that the tests in both listening and speaking appear to be valid measures of the prioritized objectives selected for this study. It should be noted, however, that testing in these areas is a relatively new development. In fact, to our knowledge, a criterion-referenced test in these areas does not exist. Because of this, our conclusions and recommendations may appear to be somewhat more tenuous than would be those of test makers in the more developed areas. For this reason, and because the tests were intended as Field Tests only and therefore were used on a limited sample, a number of caveats are stated below:

A. The objectives used as a basis for item design were those selected and prioritized by the Research Team and the Steering Committee. (See the Preface and Chapter 2 of the Technical Report for a description of the objectives development process.) Thus they do not represent the total scope of objectives that one might wish to include and assess in the listening and speaking areas of the language arts/English programs in elementary and secondary schools.

B. In the Oral Speaking Test, some of the sub-scales appear to account for substantial portions of the variance. Specifically, Voice Production, Linguistic Factors (Words) and Organization of Details correlate highly with Speaking Total. (It is possible to suggest that some sub-scales could be dropped. This would have the effect, however, of omitting items which measure high-priority objectives, and we do not feel that it is wise to make any deletion of items at this time.) It may also be said that the Written Speaking Test contributes little to the variance. It is our feeling, however, that students need to perform the Written Speaking Test in order to prepare adequately for actual production. Furthermore, this portion of the battery measures a distinct ability which relates clearly to at least some high-priority objectives not measured elsewhere.

C. Estimates of reliability were based on inter-item data analysis only. This, in effect, is a measure of homogeneity of items. No data were available which were based on parallel forms which could provide test-retest reliability estimates.

D. Rater reliability in the Oral Speaking Test was based on Pilot Test data. While there was some difficulty in establishing an adequate level of rater reliability, this had more to do with the quality of the tapes than with the teachers' rating abilities. An alternative might have been to have several raters in the room where the students were responding. However, this would probably have created a very artificial condition and would have interferred with test results. (Rooms with one-way window facilities might solve this problem.)

E. One of the areas with which we were concerned and which proved to be extremely difficult was that of devising items which would show developmental trends from grade to grade. In fact, a number of items were designed which "crossed" grade levels. They appeared to have face and content validity, but statistical analysis did not yield sufficient information to indicate clearly that they gave valid evidence of developmental trends.

Recommendation 1-1

We recommend that the Speaking and Listening Final Tests be given to larger samples of the Alberta school population, and that further analyses be performed on the results with a view of establishing acceptable performance criteria.

Recommendation 1-2

Alberta Education should determine appropriate techniques for ensuring that data derived from future testing programs in listening and speaking are carefully analyzed, collated, and "banked" so that useful comparisons can be made over time. It would also be useful if further testing in all language arts areas were performed on the same sample. This would make further correlational analysis possible.

Recommendation 1-3

In an effort to establish developmental trends, it may be useful to administer each Listening and Speaking Test at one grade level above and below the level for which the test is designed. Thus the Grade 3 test would be given to Grades 2 and 4, the Grade 6 test to Grades 5 and 7, etc. This would, of course, be possible only in the downward direction for the Grade 12 tests. Such test administration would also prove useful in the development of performance criteria at all grade levels.

Recommendation 1-4

No possibility for shortening the Listening or Speaking Tests appears appropriate. Since these tests are based on stimulus sub-scales, and since these sub-scales contain a mixture of items/objectives measures, they must be given in intact form. Thus we recommend that the entire Listening Test and the entire Speaking Test for each grade level be given as presented. (However, each test in its entirety need not be administered in one sitting.) Additionally, future study is required to determine if preparation for oral production makes a real difference in the way students actually perform.

Future Testing and Data Analysis

Conclusion 2: Statistical analysis of data largely revolves around judgments made by the researchers. Many reasonable arguments could be made for using techniques other than those reported in this study. For example, we would draw attention to the correlational analysis in which sub-scale correlations were based on data derived from item measures structured on the basis of a-priori domains. If these domains were not originally correctly conceived, this analysis might not be adequate. The factor analytic technique may also be questioned. We felt that the use of the Pearson Product-Moment Correlation Coefficients was acceptable, given the nature of the data. Tetrachoric correlations might, however, be more acceptable. Finally, basing discrimination indices on stimulus sub-scales may also be questionable. However, this appeared appropriate for the purpose of field testing, particularly in view of the fact that criterion-referenced tests were being constructed to which traditional analyses did not necessarily apply.

Recommendation 2-1

Further test reliability data should be collected. Since it is extremely difficult to construct parallel forms, a sample of students at each grade level should be selected and the tests should be given at two different points in time so that test-retest reliability can be established.

Recommendation 2-2

More data are needed to establish rater reliability. We recommend that this be dealt with as a separate study in which raters observe students from some area where they cannot be seen by the students taking the tests.

Recommendation 2-3

Continuing efforts should be made to devise items which will reveal developmental trends in the areas of listening and speaking. It is quite likely that this may need to involve the development of additional and perhaps even more specifically stated objectives.

Recommendation 2-4

Further statistical analyses should be done when data are derived from the Final Test samples. It is also suggested that analyses involving correlational techniques be used in order to devise sub-scales other than those based on a-priori domains.

Recommendation 2-5

Further factor analytic analyses should be undertaken based on the larger Final Test (and/or other) samples and should use other forms of correlation--for example, tetrachorics. Such samples would also permit the study of special sub-groups--that is, male/female; rural/urban; high/low abilities; etc.

Recommendation 2-6

When larger samples have been tested, it should be possible to perform analyses more suited to criterion-referenced testing. While these methods of analysis (item curve characteristics, etc.) are in the neonate stage, some are becoming sufficiently well explicated to be of use. At the same time, it is our feeling that traditional analysis techniques will still be useful in providing full and complete data upon which to base appropriate conclusions.

Future Test Development

Conclusion 3: Since the tests developed in this study appear to be viable, it becomes necessary to look at possible further test development. The tests constructed for this study are limited necessarily to the measurement of those objectives which formed the measurement of those objectives which formed the basis of their design. Thus they cannot reflect all of the current listening and speaking aspects of the English/language arts curriculum, nor do they purport to reflect what might reasonably be included in such a curriculum. Nevertheless, the items were based on objectives which received high priority in the literature as well as on those which were perceived by competent judges and experts to be valid and important.

Recommendation 3-1

While objectives for this study were carefully selected, further energy should be expended in continuous review and upgrading of objectives. A mechanism is needed for ensuring that such objectives become a part of the language arts/English curriculum. Such ongoing development would also require constant review and development measurement techniques.

Recommendation 3-2

The Research Team feels that the Listening Comprehension Tests developed in England by Wilkinson, Stratta, and Dudley (1974) would be a most useful addition to a Listening Test battery for Alberta schools. While these tests measure objectives not specifically dealt with in this study for reasons described earlier in this report, they are based on objectives considered highly relevant and valid by experts in the language arts field. If this test is to be used, Canadianization is essential.

Recommendation 3-3

The tests provided in this study are useful as initial tools for identifying students who have difficulties in listening and speaking. However, the tests were not primarily designed as diagnostic tests of listening/speaking difficulties, and therefore other tests suitable for such purposes need to be developed and/or selected, and made available to teachers.

Attitude Testing

Conclusion 4: The data in this study indicate that there is no relationship between the attitude measure and any of the other measures used in this study. (This was a rather surprising finding because it seemed reasonable to expect that a measure of communication anxiety should correlate with the area of speaking, at least.) A number of reasons for this lack of relationship have been suggested in the Final Report. Since the measure of attitude used was a very "pure" one--that is, it loaded on only one factor--it may be that a correlation of any magnitude could not reasonably have been expected. In effect, then, we can conclude that the Attitude Tests used may very well measure communication anxiety, but that such anxiety does not necessarily affect a person's ability to perform the tasks required in the Speaking and Listening Tests as structured in this study.

Recommendation 4-1

We do not recommend the use of the *Personal Report of Communication Fear* or the *Personal Report of Communication Anxiety* (McCrosky, 1977) as part of the testing program in listening and speaking. These tests may well be useful as measures of communication anxiety, but their use should be limited to measuring only this aspect of student development.

Recommendation 4-2

If MACOSA deems it necessary to provide attitude tests for the listening/speaking areas of the language arts curriculum, a specific program for developing such tests should be undertaken. It is our opinion that tests of attitudes toward school, such as the ones currently being developed for MACOSA, might well be feasible.

Reporting Test Results

Conclusion 5: One of the concerns in any province-wide testing program relates to dissemination and security. In our opinion, it would not be wise to release any of the Pilot, Field, or Final Listening or Speaking Test items or stimulus materials. These are, after all, intended for assessment rather than instructional purposes.

Recommendation 5-1

No tests or stimulus materials should be released, and rigid control should be exercised over copyright, test administration, return of copies of the tests tapes, etc.

Recommendation 5-2

The data collected as a result of the administration of the

Final Tests should not be disseminated to participating schools or other interested parties until acceptable performance criteria have been established. Release of such data should be carefully controlled, and clear and precise explanatory material should be provided indicating the purposes, functions, and limitations of the test instruments.

Recommendation 5-3

The Research Team recommend that the tests submitted with the Final Report be administered to a large sample of Alberta students as indicated in the recommendations outlined above. As a result of such a testing program, performance criteria could be established and the tests could be used as a criterion measure whenever Alberta Education or a school system requests such an evaluation.

Recommendation 5-4

Alberta Education should prepare a statement of those objectives it thinks are appropriate as part of the school program in listening and speaking at the elementary and secondary grade levels. Such a list should indicate which objectives are measured by the Final Tests prepared for this study. This statement should be made available to school systems throughout Alberta for use in future curriculum planning.

Teacher Training

Conclusion 6: Since the teaching of listening and speaking has not been emphasized until recently in the Alberta school language arts/English curriculum, and since it has not been a major element in teacher education programs, it is important to ensure that teachers are provided with the requisite skills necessary for effective initiation and development of the listening and speaking components of the school program.

Recommendation 6-1

Pre-service and in-service training at all levels should be made available to teachers concerned with providing language arts/English instruction in elementary and secondary schools, in order that they can develop and expand the requisite skills for teaching and diagnosing listening and speaking knowledges and skills as well as for integrating listening and speaking into the total language arts/English curriculum.

Recommendation 6-2

Pre-service and in-service training in the area of functional communication skills should be made available to all teachers, regardless of their curriculum speciality.

Recommendation 6-3

When the Final Listening and Speaking Tests are in general use in the province of Alberta, pre-service and in-service training should be provided to teachers in both the administration of the tests and the interpretation of test results. Further, training should be provided to teachers in the use of the test results in curriculum planning at the individual, class, school, and system levels.

REFERENCES

Allen, R. and K. Brown (eds.). Developing Communication Competence in Children. Skokie, Illinois: National Textbook Company, 1976.

Ammon, R. "Listening as a Means of Developing Language," Elementary English, 51, April, 1974, 515-518.

Anderson, M., E. Nichols, and H. Booth. The Speaker and His Audience, Second Edition. New York: Harper and Row, 1974.

Armstrong, C. and P. D. Brandes. The Oral Interpretation of Literature. New York: McGraw-Hill, 1963.

Arndt, C. and J. Husband. "Listen," English Journal, 29, May, 1940, 375.

Bacon, W. A. The Art of Interpretation. New York: Holt, Rinehart and Winston, 1966.

Baird, A., and F. Knower. Essentials of General Speech. New York: McGraw-Hill, 1968.

Barbara, D. "On Listening--the Role of the Ear in Psychic Life," Today's Speech, 1957, 12-15.

Barker, L. Listening Behavior. Englewood Cliffs, New Jersey: Pentice-Hall, 1971.

Barnes, D. and J. Britton. Language, the Learner, and the School. Baltimore, Maryland: Penguin Books, 1969.

Berlo, David K. The Process of Communication. New York: Holt, Rinehart and Winston, 1960.

Bloom, B. S. (ed.). A Taxonomy of Educational Objectives, Handbook I, The Cognitive Domain. New York: David McKay, 1956.

Book, C. and K. Galvin. Instruction in and about Small Group Instruction. Urbana, Illinois: ERIC Clearinghouse on Reading and Communication Skills, 1975.

Bormann, E., R. Nichols, and G. Shapiro. Interpersonal Communication in the Modern Organization. Englewood Cliffs, New Jersey: Prentice-Hall, 1969.

Britton, J. (ed.). Talking and Writing: A Handbook for English Teachers. London: Methuen, 1967.

Brown, C. and C. Van Riper. Speech and Man. Englewood Cliffs, New Jersey: Prentice-Hall, 1966.

Brown, J. and G. R. Carlsen. Brown-Carlsen Listening Comprehension Test. New York: Harcourt, Brace and World, 1955.

Brown, K. "Speech and Listening in Language Arts Textbooks, Part I," Elementary English, 44, April, 1967, 336-341.

__________. "Speech and Listening in Language Arts Textbooks, Part II." Elementary English, 44, May, 1967, 461-465.

Caffrey, J. "An Introduction to the Auding Concept." Education, 70, December, 1949, 234-239.

Callaghan, J. "Testing Ability to Organize Ideas," in Cayer, R., J. Green, and E. Baker (eds.). Listening and Speaking in the English Classroom. 1971.

Campbell, D. "Systematic Error on the Part of Human Links in Communication Systems." Information and Control, 1958, 334-369.

Cayer, R., J. Green and E. Baker (eds.). Listening and Speaking. Toronto: Macmillan, 1971.

Cazden, C. Language and Education. New York: Holt, Rinehart and Winston, 1972.

Clark, M., E. Erway, and L. Beltzer. The Learning Encounter: The Classroom as a Communications Workshop. New York: Random House, 1971.

Devine, T. "Listening: What Do We Know after Fifty Years of Research and Theorizing?" Journal of Reading, 21, January, 1978, 296-304.

DeVito, J. The Psychology of Speech and Language: An Introduction to Psycholingusitics. New York: Random House, 1970.

Douglas, J. "The Measurement of Speech in the Classroom." Speech Teacher, November, 1958, 31-33.

Ecroyd, D. Speech in the Classroom. Englewood Cliffs, New Jersey: Prentice-Hall, 1960.

Egan, G. Encounter: Group Process for Interpersonal Growth. Belmont, California: Brooks/Cole, 1970.

Fawcett, A. "Training in Listening." Elementary English, 43, May, 1966, 473-476.

Fernandez, T. L. (ed.). Oral Interpretation and the Teaching of English. Illinois: National Council of Teachers of English, 1969.

Fessenden, S., R. Johnson, P. Larson, and K. Good. Speech for the Creative Teacher. Dubuque, Iowa: William C. Brown, 1968.

Fisher, I. "Reluctant Listeners." The Speech Teacher, 22, 1973, 54-57.

Furth, H. Piaget for Teachers. Englewood Cliffs, New Jersey: Prentice-Hall, 1970.

Galloway, C. Teaching is Communicating: Nonverbal Language in the Classroom. Washington, D. C.: National Education Association, 1970.

Galvin, K. and C. Book. Speech Communication: An Interpersonal Approach for Teachers. Skokie, Illinois: National Textbook Company, 1972.

Geiger, D. The Sound, Sense, and Performance of Literature. Chicago, Illinois: Scott, Foresman, 1963.

Goldberg, A. and C. Larson. Group Communication: Discussion Processes and Applications. Englewood Cliffs, New Jersey: Prentice-Hall, 1975.

Gray, J. W. (ed.). Perspectives on Oral Interpretation. Auburn, Alabama: Burgess Publishing Co., 1968.

Greene, H. A. and W. T. Petty. Developing Language Skills in the Elementary School. Boston: Allyn and Bacon, 1959.

Gruner, C. "Behavioral Objectives for the Grading of Classroom Speeches." Speech Teacher, 25, September, 1968, 27-30.

Halliday, M. Explorations in the Functions of Language. London: Edwin Arnold, 1974.

Horner, H. Investigation of the R-Mode of Communication: Reading and Listening. Master's Thesis, University of Denver, 1973.

Hollow, Sr. M. *An Experimental Study of Listening at the Intermediate Grade Level*. Unpublished Doctoral Dissertation, Fordham University, 1955.

Holtzman, P. and D. Ecroyd. *Communication Concepts and Models*. Skokie, Illinois: National Textbook Company, 1976.

Hooper, R. and R. Naremore. *Children's Speech: A Practical Introduction to Communication Development*. New York: Harper and Row, 1973.

Horrworth, Gloria. "Listening: A Facet of Oral Language." *Elementary English*, 53, December, 1966, 856-864.

Huckleberry, A. and E. Strother. *Speech Education for the Elementary Teacher*. Boston: Allyn and Bacon, 1966.

Hurt, H., M. Scott, and J. McCroskey. *Communication in the Classroom*. Reading, Massachusetts: Addison-Wesley, 1978.

Inhelder, R. and J. Piaget. *The Early Growth of Logic in the Child*. New York: W. W. Norton, 1969.

Johnson, W., F. L. Darley and D. C. Spriestersbach. *Diagnotic Methods in Speech Pathology*. New York: Harper and Row, 1963.

Johnson, W. *People in Quandries*. New York: Harper and Row, 1946.

Kelley, W. "Objectivity in the Grading and Evaluation of Speeches," in R. Cayer, J. Green, and E. Baker (eds.). *Listening and Speaking*. Toronto: Macmillan, 1971.

Keltner, J. *Interpersonal Speech Communication: Elements and Structures*. Belmont, California: Wadsworth, 1970.

Kibler, R. and L. Barker. "Stating Objectives in Speech Communication." *Bulletin of the National Association of Secondary School Principals*, 54, December, 1970, 32.

Klatt, I. M. *Language Processing Skills of Poor Readers and Poor Articulators in Grade Three*. Unpublished Master's Thesis, University of Calgary, 1972.

Klein, M. *Talk in the Classroom*. Urbana, Illinois: ERIC Clearinghouse on Reading and Communication Skills, 1977.

Knower, F. "The Present State of Experimental Speech-Communication Research," in P. Reed (ed.). *The Frontiers in Experimental Speech Communication Research*. New York: Speech Association of America, 1966.

Landry, D. "The Neglect of Listening." Elementary English, 46, May, 1969, 599-605.

Larson, C., P. Backlund, M. Redmond, and A. Barbour. Assessing Functional Communication. Falls Church, Virginia: Speech Communication Association, 1978.

Lee, C. I. Oral Interpretation, Third Edition. Boston: Houghton-Mifflin, 1965.

Lindvall, C. M. (ed.). Defining Educational Objectives. Pittsburgh: University of Pittsburgh Press, 1964.

Loban, W. The Language of Elementary School Children. Urbana, Illinois: National Council of Teachers of English, 1963.

__________. Language Development: Kindergarten Through Grade Twelve. Urbana, Illinois: National Council of Teachers of English, 1976.

Lundsteen, S. Listening--Its Impact on Reading and the Other Language Arts. Urbana, Illinois: National Council of Teachers of English, 1971.

__________. Children Learn to Communicate. Englewood Cliffs, New Jersey: Prentice-Hall, 1976.

Marchak, N. et. al. Assessing Communication Skills: A Review of the Literature. Report to the Minister's Advisory Committee on Student Achievement (MACOSA). Alberta Education, 1979.

Mattingly, A. S. and W. H. Grimes. Interpretation, Second Edition. Belmont, California: Wadsworth Publishing, 1970.

McCroskey, J. "Oral Communication Apprehension: A Summary of Recent Theory and Research." Human Communication Research, 4, Fall, 1977, pp. 78-96.

McCroskey, J. and J. Daly. "Teacher's Expectations of the Communication Apprehensive Child in the Elementary School." Human Communication Research, 3, Fall, 1976, pp. 67-72.

Mead, N. Issues Relating to Assessing Listening Ability. Unpublished Paper Presented at the Annual Conference of the American Educational Research Association, Toronto, 1978.

Miller, G. "Communication Process: An Operational Approach." *Ohio University for Communication Studies*, 16, November, 1967, 2-8.

Mudd, C. and M. Sillars. *Speech Content and Communication*. Scranton, Pennsylvania: Chandler, 1969.

Mulgrave, D. *Speech for the Classroom Teacher, Third Edition*. Englewood Cliffs, New Jersey: Prentice-Hall, 1961.

Parrish, W. M. *Reading Aloud, Fourth Edition*. New York: Ronald Press, 1966.

Phillips, G. "A New Direction for the Speech Profession," in J. Akin, A. Goldberg, G. Meyers, and Steward (eds.). *Language Behavior*. 19, The Hague: Mouton, 1970.

Plattor, E. *Assessing Listening Skills--An Examination of Selected Research Reports*. Unpublished Report to the Minister's Advisory Committee on Student Achievement (MACOSA). Edmonton, Alberta: Alberta Education, 1977.

Popham, W. J. *Criterion-Referenced Measurement*. Englewood Cliffs, New Jersey: Prentice-Hall, 1978.

Purves, A. et. al. *Common Sense and Testing in English*. Urbana, Illinois: National Council of Teachers of English, 1975.

Reid, L. *Teaching Speech, Fourth Edition*. New York: McGraw-Hill, 1971.

Rosenfeld, L. *Human Interaction in the Small Group Setting*. Columbus, Illinois: Charles Merrill, 1973.

Sanson, C. *Speech and Communication in the Primary School*. London: A. and C. Black, 1965.

Sommer, R. *Personal Space*. Englewood Cliffs, New Jersey: Prentice-Hall, 1969.

Streeter, D. *A Speech Handbook for Teachers*. Englewood Cliffs, New Jersey: Prentice-Hall, 1964.

Taylor, S. *Listening*. Washington, D. C.: National Education Association, 1973.

Tutolo, D. "A Cognitive Approach to Teaching Listening." *Language Arts*, 54, March, 1977, 262-265.

Vogelsang, R. et. al. *Speech Arts in the Elementary School*. Portland, Oregon: Portland State University, 1972.

Vygotsky, L. *Thought and Language*. Cambridge, Massachusetts: M.I.T. Press, 1962.

Weaver, A., G. Borchers, and D. Smith. *The Teaching of Speech*. Englewood Cliffs, New Jersey: Prentice-Hall, 1963.

Weaver, Carl. *Human Listening*. New York: Bobbs-Merrill, 1972.

Weaver, S. and W. Rutherford. "A Hierarchy of Listening Skills." *Elementary English*, 51, November/December, 1974.

Wheeless, L. "An Investigation of Receiver Apprehension and Social Context Dimensions of Communication Apprehension." *Speech Teacher*, 24, 1975, 261-268.

White, B. "Critical Influences in the Origins of Competence." *Merrill-Palmer Quarterly*, 21, 1975, 243-266.

Wiemann, M. and J. Wiemann. *Nonverbal Communication in the Elementary Classroom*. Urbana, Illinois: ERIC Clearinghouse on Reading and Communication Skills, 1975.

Wilkinson, A. "The Concept of Oracy." *English Journal*, 59, January, 1970.

Wilkinson, A., L. Stratta, and P. Dudley. *The Quality of Listening*. Toronto: Macmillan of Canada, 1974.

__________. *The Listening Comprehension Tests*. Toronto: Macmillan of Canada, 1974.

Wood, B. (ed.). *Development of Functional Communication Competencies: Grades 7-12*. Urbana, Illinois: ERIC Clearinghouse on Reading and Communication Skills, 1977.

APPENDICES

Appendix A-1

Initial Listening/Speaking Test Objectives: Knowledges[1]

Listening/Speaking Test Objectives	Grade Level Priority[2]			
	3	6	9	12
Knowledges (The student will know . . .)				
1. Environmental Sound Factors				
1. the relationship of environmental sounds to their environmental referrents.	1	1	1	1
2. the effects of environmental sounds on feelings, emotions, and attitudes.	1	1	1	1
2. Voice Production Factors				
1. the ways in which voice production factors (volume, pitch, juncture, quality, rate and fluency) and the variations of these factors convey the meaning and/or mood of a message and maintain audience interest.	2	2	2	2
2. the ways in which voice production factors and the variation of these factors affect the aesthetic quality of a communication.	1	1	1	1
3. the points in a communication which should be emphasized to facilitate meaning, mood, and/or audience interest.	1	3	3	2
3. Nonverbal Factors				
1. the ways in which nonverbal factors (posture, gesture and body movement, audience contact, and facial expression) and the variation of these factors convey the meaning and/or mood of a message and maintain audience interest.	1	1	2	2
2. the ways in which nonverbal factors and the variations of these factors affect the aesthetic quality of a communication.	1	1	2	2
3. the points in a communication where specific nonverbal factors should be used to facilitate meaning, mood, and/or interest.	1	1	3	3

---

[1] Knowledges listed in this section are associated with oral rather than written language, except when the student is listening to a prepared piece of writing (e.g.: oral reading).

[2] 3 = high priority; 2 = some priority; 1 = minimal priority.

Listening/Speaking Test Objectives (Continued)	Grade Level Priority 3	6	9	12
Knowledges (The student will know . . .)				
4. Linguistic Factors				
4.1 General Linguistic Factors				
1. the ways in which linguistic factors (sounds, words, and sentences) and the variations of these factors affect the meaning and/or mood of a message and maintain audience interest.	1	1	2	2
2. the ways in which linguistic factors and the variations of these factors affect the aesthetic quality of a communication.	1	1	1	1
3. what fillers (um, ah, uh, etc.), redundancies, and stabilizers (you know, you see, eh, etc.) are.	1	2	3	3
4. the ways in which fillers, redundancies, and stabilizers act as facilitators or barriers in communication.	1	1	2	2
5. the ways in which linguistic constraints (predictive devices) are used in facilitating meaning and/or mood and maintaining audience interest.	2	2	2	2
6. the sound, word, and sentence features appropriate for use with different language levels.[1]	1	2	3	3
4.2 Specific Linguistic Factors				
A. Sounds				
1. conventional speech sounds and sound clusters.	3	3	1	1

---

[1] Language levels may be classified as casual, informal, and formal. Each level has specific features and is used for a specific audience, purpose, situation, and format in an oral communication interaction. It should be noted that these features vary in oral and written communication.

Listening/Speaking Test Objectives (Continued)	Grade Level Priority			
	3	6	9	12
Knowledges (The student will know . . .)				
4. 4.2 A. 2. dialectic speech sounds and sound clusters.	1	1	1	1
3. phonetic and structural generalizations and their applications to word pronunciation.	3	3	2	1
4. conventional pronunciations of parts of words, and words.	3	3	2	1
5. dialectic pronunciations of parts of words, and words.	1	1	1	1
6. positions of sounds, clusters of sounds, and sequences of sounds in words.	3	3	1	1
B. Words (Meanings and Structure)				
1. an oral receptive and expressive vocabulary consistent with the age norm.	3	3	3	3
2. conventional word forms (e.g.: noun forms, verb forms, etc.) consistent with the age norm.	3	3	3	1
3. dialectic word forms (for use in dialogue, monologue, etc.).	1	1	1	1
4. a variety of word building techniques to meet the requirements of clarity, precision, extensiveness, and stylistic appropriateness.	1	2	2	2
5. a variety of transitional devices.	2	2	2	2
6. connotative and denotative meanings of words and expressions in context.	1	3	3	3
7. the impact of connotative words as used in specific contexts.	1	1	1	1
C. Sentences				
1. conventional sentence patterns, levels, and types to facilitate meaning, mood, and/or interest.	3	3	3	3

Listening/Speaking Test Objectives (Continued)	Grade Level Priority 3	6	9	12
Knowledges (The student will know . . .)				
4. 4.2 C. 2. a variety of sentence building techniques to meet the requirements of clarity, precision, extensiveness, and stylistic appropriateness.	1	2	3	3
3. a variety of transitional devices for clarity and emphasis.	2	2	2	2
4. a variety of stylistic devices to facilitate meaning, mood, and/or audience interest, (e.g.: repetition of sentence parts, repetition of words, parallelism, etc.).	1	1	3	3
5. the ways in which stylistic devices affect meaning, mood, and/or audience interest (e.g.: Repetition of sentence parts, repetition of words, parallelism, etc.).	1	1	2	2
5. Organizational Factors				
A. Ideational Fluency				
1. a variety of associations in the form of ideas, opinions, feelings, and attitudes in terms of a given stimulus.	3	3	3	3
2. what questions to ask about a stimulus in order to make and extend associations.	1	2	2	2
B. Plan Analysis				
1. Selecting a topic.	3	3	3	3
2. techniques involved in narrowing a topic in terms of preparation and presentation time.	1	2	3	3
3. techniques involved in collecting and recording information.	1	1	1	1
4. techniques involved in developing point of view (mental/physical; first/third person, etc.) and maintaining point of view consistently.	1	2	3	3

Listening/Speaking Test Objectives (Continued)	Grade Level Priority			
	3	6	9	12
Knowledges (The student will know . . .)				
5. B. 5. requirements of different roles and techniques involved in selecting an appropriate speaking role, and maintaining it consistently.	1	1	1	1
6. techniques involved in selecting a purpose.	3	3	3	3
7. techniques involved in creating and maintaining a variety of audience effects	2	2	2	2
8. features of various language levels and techniques for determining appropriate language level selection.	2	2	2	2
C. Organization of Details				
1. what a principal (main) idea is and what subordinate ideas are.	2	2	2	2
2. techniques involved in developing principal and subordinate ideas.	1	2	3	3
3. techniques involved in appropriate placement of subordinate and main ideas.	1	2	3	3
4. techniques involved in selecting appropriate details to support main and subordinate ideas.	2	2	2	2
5. techniques involved in developing and placing topic sentences.	1	1	2	2
6. various kinds of supporting details and techniques involved in appropriate selection.	1	2	2	2
7. various kinds of supporting materials and techniques involved in their appropriate selection.	1	2	2	2
8. techniques involved in developing beginnings, middles, and endings.	3	3	3	3
9. organizing patterns used in sequencing details and techniques involved in selection of organizing patterns.	1	2	2	2

Listening/Speaking Test Objectives (Continued)	Grade Level Priority			
	3	6	9	12
Knowledges (The student will know . . .)				
5. C. 10. techniques involved in sequencing details according to a specified organizing pattern.	2	2	2	2
11. types of transitional devices and techniques involved in appropriate selection and use.	1	2	3	3
12. techniques involved in message emphasis (e.g.: repetition, restatement, etc.).	1	1	2	3
13. techniques involved in developing characterization.	1	1	1	1
14. techniques involved in developing setting.	1	1	1	1
15. techniques involved in developing plot.	1	1	1	1
16. techniques involved in developing theme.	1	1	1	1
17. amount and type of information required for effective inferencing and critical evaluation.	1	2	3	3
18. techniques involved in summarizing and paraphrasing.	2	2	2	2
19. techniques involved in developing an outline.	1	2	2	2

6. Evaluation

A. judge the impact of environmental sound factors. (See #1 for objectives and priorities).

B. evaluate a communication in terms of voice production factors. (See #2 for objectives and priorities).

C. evaluate a communication in terms of nonverbal factors. (See #3 for objectives and priorities).

D. evaluate a communication in terms of linguistic factors. (See #4 for objectives and priorities).

E. evaluate a communication in terms of organizational factors. (See #5 for objectives and priorities).

Appendix A-2

Initial Listening Test Objectives: Skills[1]

Listening Test Objectives	Grade Level Priority[2]			
	3	6	9	12
Skills[1] (The student will be able to . . .)				
1. Environmental Sound Factors				
A. Identification				
1. match environmental sounds with their environmental referrents.	3	1	1	1
2. distinguish sounds that are appropriate within a context from those which are inappropriate and/or distracting.	3	2	2	2
B. Association				
1. produce associations to sounds with their conventional meanings.	3	1	1	1
2. extend associations to sounds beyond their conventional meanings.	1	2	1	1
3. match environmental sounds with associated feelings, emotions, and/or attitudes.	3	2	2	2
2. Voice Production Factors				
A. Volume				
1. relate vocal stress to meaning, mood, and/or interest.	2	2	2	2
B. Pitch				
1. relate intonation to meaning, mood, and/or interest.	2	2	2	2
C. Juncture				
1. relate pauses between syllables, words, phrases, and sentences to meaning, mood, and/or interest.	2	2	2	2

---

[1] Skills listed in this section are associated with oral rather than written language, except when the student is listening to a prepared piece of writing (e.g.: oral reading).

[2] 3 = high priority; 2 = some priority; 1 = minimal priority.

Listening Test Objectives (Continued)	Grade Level Priority			
	3	6	9	12
Skills (The student will be able to . . .)				
2. D. Quality				
1. relate vocal quality to meaning, mood, and/or interest.	3	3	3	3
E. Rate and Fluency				
1. relate pace to meaning, mood, and/or interest.	3	3	3	3
2. relate variations in pausing and phrasing to meaning, mood, and/or interest.	2	2	2	2
3. maintain interest in spite of the use of non-facilitating fillers (um, ah, uh, etc.) repetitions of words and phrases, and/or stabilizers (you know, you see, eh, etc.).	1	1	1	1
3. Nonverbal Factors				
A. Posture				
1. relate posture to meaning, mood, and/or message.	1	1	1	1
B. Gesture and Body Movement				
1. relate gestures to meaning, mood, and/or interest.	1	1	1	1
2. relate body movements to meaning, mood, and/or interest.	1	1	1	1
C. Audience Contact				
1. relate eye contact to meaning, mood, and/or message.	1	1	1	1
2. relate physical distance to meaning, mood, and/or message.	1	1	1	1
D. Facial Expression				
1. relate facial expressions to meaning, mood, and/or interest.	1	1	1	1

Listening Test Objectives (Continued)	Grade Level Priority			
	3	6	9	12
Skills (The student will be able to . . .)				
4. Linguistic Factors				
A. Sounds (Articulation and Pronunciation)				
1. discriminate speech sounds.	3	1	1	1
2. discriminate conventional[1] pronunciation of speech sounds clusters of sounds, and sequences of speech sounds which comprise meaning units.	3	3	1	1
3. discriminate dialectic[2] pronunciation of speech sounds clusters of sounds, and sequences of speech sounds which comprise meaning units.	1	1	1	1
4. relate dialectic pronunciation to meaning, mood, and/or interest (e.g.: in dialogue, monologue, etc.)	1	1	1	1
5. repeat sequences of sound, sound clusters, and sound sequences which comprise meaning units in a designated order (includes repetition of words and sentences).	3	2	1	1
6. identify the position of sounds and sound clusters (vowels, consonants, digraphs, etc.) in words.	1	1	1	1
7. use phonetic and structural generalizations to discriminate the pronunciation of unknown words.	1	1	1	1
8. supply missing sequences of sounds in meaning units within specific contexts (Phonological constraints).	2	2	2	2

---

[1] Conventional is defined here as the pronunciation agreed upon by people who live in the same society. It is well-known and well-accepted pronunciation.

[2] Dialectic is defined here as the variant pronunciation caused by regional and individual speech patterns.

Listening Test Objectives (Continued)	Grade Level Priority			
	3	6	9	12
Skills (The student will be able to . . .)				
4. B. Words (Meaning and Structure)				
1. understand the meaning of denotative words in specific contexts.	3	3	3	3
2. understand the meaning of dialectic words.	1	1	1	1
3. understand connotative meanings of words in specific contexts.	1	2	3	3
4. understand the impact of connotation in specific contexts.	1	1	1	1
5. understand the meaning of figurative words and expressions in specific contexts.	2	2	2	2
6. supply missing words appropriately within specific contexts (semantic constraints).	2	2	2	2
7. understand a variety of types of words as they are used to facilitate meaning, mood, and/or interest (e.g.: vivid words and expressions, precise words, etc.).	3	3	3	3
8. identify the types of words appropriate for use with the selected language level.[1]	1	1	1	1
C. Sentences (Meaning and Structure)				
1. understand messages presented in various sentence (grammatical) patterns (s-v-o, etc.).	3	2	1	1

---

[1] Language levels may be classified as casual, informal, and formal. Each level has specific features and is used for a specific audience, purpose, situation and format in an oral communication interaction. It should be noted that these features vary in oral and written communication.

Listening Test Objectives (Continued)	Grade Level Priority			
	3	6	9	12
Skills (The student will be able to . . .)				
4. C. 2. understand messages presented at various sentence levels (complete/incomplete, simple/compound/complex, etc.).	2	2	3	3
3. understand messages presented in various sentence types (statements, questions, etc.).	3	3	2	1
4. understand transitional devices used to clarify and emphasize meaning and mood.	1	2	3	3
5. identified the purpose to stylistic devices (e.g.: Repetition of sentence parts, repetition of words, parallelism, etc.) to facilitate meaning, mood, and/or interest.	1	1	3	3
6. identify the sentence patterns, levels and types appropriate for use with the selected language level.	1	1	1	1
7. supply missing parts of sentences appropriately within specific contexts (syntactic constraints).	2	2	2	2
5. Organizational Factors				
A. Ideational Fluency				
1. provide associations in the form of ideas, opinions, feelings, and attitudes produced in terms of a given stimulus.	1	1	1	1
2. identify associations in the form of ideas, opinions, feelings, and attitudes produced in terms of a given stimulus.	1	1	1	1
3. extend associations beyong the understanding level to facilitate meaning, mood, and/or interest.	1	2	3	3
B. Plan Analysis				
1. indicate the selected topic.	3	3	3	3

Listening Test Objectives (Continued)	Grade Level Priority			
	3	6	9	12
Skills (The student will be able to . . .)				
5. B. 2. indicate the speaker's purpose in terms of audience and situation.	1	2	3	3
3. indicate the speaker's role.	1	1	2	3
4. indicate the speaker's point of view.	1	1	1	1
5. indicate the selected language level.	1	2	3	3
C. Comprehension of Organized Details				
1. Literal Level[1]				
A. indicate the main idea of a set of details	3	3	3	3
B. indicate the placement of the topic sentence.	1	1	1	2
C. indicate the details that directly support the main idea.	2	2	3	3
D. indicate the different types of details that support the main idea (e.g.: statistics, expert opinions, illustrations, anecdotes, examples, analogies, etc.).	1	1	2	3
E. identify the beginning, middle, and ending (introduction, body and conclusion).	3	3	3	3
F. identify the technique used for developing the beginning, middle and ending (introduction, body and conclusion).	1	1	1	2
G. identify transitional devices.	1	2	3	3
H. identify the type of organizing pattern used and the details used to produce the pattern (e.g.: time, place, classification, etc.).	3	3	3	3

---

[1] Literal level is defined here as the memory or recognition/recall level.

Listening Test Objectives (Continued)	Grade Level Priority			
	3	6	9	12
Skills (The student will be able to . . .)				
5. C. 1. I. indicate the details used to develop setting.	1	2	2	2
J. indicate the details used to develop characterization.	1	2	2	2
K. indicate the details used to develop plot (conflicts, complications, suspense, climax, outcome).	1	2	2	2
L. indicate the theme.	1	1	1	2
M. provide a summary.	2	2	2	2
N. provide a paraphrase.	1	1	2	2
2. Interpretational Level[1]				
A. indicate the reason for the placement of the main idea statement.	1	1	2	3
B. elicit additional related and appropriate information where the presented details appear inadequate for interpretation.	2	2	2	2
C. infer thoughts, feelings, emotions of others (including the speaker) in terms of the communication.	3	3	3	3
D. draw conclusions in terms of the intent of the communication.	2	2	2	2
E. anticipate alternative consequences and outcomes.	2	2	2	2
F. suggest alternative hypotheses and courses of action.	3	3	3	3

---

[1] Interpretational level is defined here as the relationship level (e.g.: seeing similarities and differences, grouping within and among facts, making generalizations, etc.). It includes the application, analysis, and synthesis phases of listening.

Listening Test Objectives (Continued)		Grade Level Priority 3	6	9	12
Skills (The student will be able to . . .)					
5. C. 2. G.	distinguish between stated and implied details.	1	1	2	2
H.	infer speaker's purpose in terms of the communication.	3	3	3	3
I.	infer speaker's conception of audience characteristics in terms of the language level used.	1	1	2	2
J.	infer speaker's point of view in terms of the communication.	1	2	2	2
K.	infer speaker's perceived role in terms of the communication.	1	1	2	2
L.	infer character traits, relationships, motivations, etc.	1	2	2	2
M.	infer sensory impressions.	2	2	2	2
N.	infer conflicts in attitudes, feelings, opinions, ideas, etc.	1	2	2	2
O.	infer mood and tone.	3	3	3	3
P.	infer theme where not stated.	1	1	2	2
3.	Critical Level[1]				
A.	judge whether the speaker's purpose was met.	3	3	3	3
B.	judge the effect of speaker's choice of role, point of view, and language level, and the effect of these on the communication.	1	1	2	3
C.	judge the potential effectiveness of the message on other audiences.	1	1	1	1

---

[1] Critical level is defined here as the valuing or judgmental level, where opinions are formed and judgments are made based on the details presented and the results of skills applied during the literal and inferential levels.

Listening Test Objectives (Continued)		Grade Level Priority			
		3	6	9	12
Skills (The student will be able to . . .)					
5. C. 3. D.	judge if the message is factual or fictitious.	3	3	3	3
E.	judge the adequacy of supporting types of details in terms of message effect.	1	1	3	3
F.	judge the appropriateness of arrangement of details.	1	1	2	2
G.	judge the appropriateness of the placement of the main idea statement.	1	1	2	2
H.	judge relevancy of details.	1	1	3	3
I.	judge the appropriateness of sources used.	1	1	1	1
J.	evaluate speaker bias or prejudice.	2	2	3	3
K.	evaluate the impact of the use of connotative language.	1	1	1	1
L.	evaluate the impact of the use of figurative language.	1	1	1	1
M.	judge the wisdom and consequences of a character's actions in terms of the information presented.	2	2	2	2
N.	evaluate the development of setting, theme and plot.	1	1	2	2

6. Evaluation Factors[1]

A. judge the impact of environmental sound factors. (See 1 for objectives and priorities.)

---

[1] Evaluation here involves a listener judging his or her own performance and prioritization is related to that of objectives listed earlier.

Listening Test Objectives (Continued)	Grade Level Priority 3	6	9	12
Skills (The student will be able to . . .)				
6. B. evaluate a communication in terms of voice production factors. (See 2 A-E for objectives and priorities.)				
C. evaluate a communication in terms of nonverbal factors. (See 3 A-D for objectives and priorities.)				
D. evaluate a communication in terms of linguistic factors. (See 4 A-C for objectives and priorities.)				
E. evaluate a communication in terms of organizational factors. (See 5 A-C for objectives and priorities.)				

Appendix A-3

Initial Speaking Test Objectives: Skills[1]

Speaking Test Objectives	Grade Level Priority[2] 3	6	9	12
Skills (The student will be able to . . .)				
1. Environmental Sound Factors				
A. use sounds to facilitate meaning, mood, and/or interest.	1	1	1	1
2. Voice Production Factors				
A. Volume				
1. project the voice appropriately (loud/soft).	3	3	3	3
2. use vocal stress to facilitate meaning, mood, and/or interest.	3	3	3	3
B. Pitch				
1. use intonation (high/low) to facilitate meaning, mood, and/or interest.	3	3	3	3
C. Juncture				
1. use pauses between syllables, words, phrases, and sentences to facilitate meaning, mood, and/or interest.	3	3	3	3
D. Quality				
1. use a pleasant voice to facilitate meaning, mood, and/or interest.	3	3	3	3
2. vary vocal quality where appropriate to facilitate meaning, mood, and/or interest (e.g.: in story telling, oral reading, etc.).	3	3	3	3

---

[1] Skills listed in this section are associated with oral rather than written language.

[2] 3 = high priority; 2 = some priority; 1 = minimal priority.

Speaking Test Objectives (Continued)	Grade Level Priority			
	3	6	9	12
Skills (The student will be able to . . .)				
2. E. Rate and Fluency				
1. use pace to facilitate meaning, mood, and/or interest.	3	3	3	3
2. vary pause and phrasing to facilitate meaning, mood, and/or interest (e.g.: in story telling, oral reading, etc.).	3	3	3	3
3. use a minimum number of non-facilitating fillers (um, ah, uh), non-facilitating repetitions of words and phrases (redundancies), and/or non-facilitating stabilizers (you know, you see, eh).	3	3	3	3
3. Nonverbal Factors				
A. Posture				
1. use posture which is consistent with the meaning and/or mood of a message.	1	1	1	1
B. Gesture and Body Movement				
1. use appropriate gestures to facilitate meaning, mood, and/or interest.	1	1	2	2
2. use appropriate body movements to facilitate meaning, mood, and/or interest.	1	1	2	2
C. Audience Contact				
1. maintain eye contact with audience.	3	3	3	3
2. maintain appropriate physical distance to audience.	1	1	1	1
D. Facial Expression				
1. use appropriate facial expressions to facilitate meaning, mood, and/or interest.	3	3	3	3
4. Linguistic Factors				
A. Sounds (Articulation and Pronunciation)				
1. articulate speech sounds clearly and precisely.	3	1	1	1

Speaking Test Objectives (Continued)	Grade Level Priority			
	3	6	9	12
Skills (The student will be able to . . .)				
4. A. 2. use conventional[1] pronunciations of speech sounds, clusters of sounds, and sequences of speech sounds in words.	3	3	1	1
3. use dialectic[2] pronunciations of speech sounds, clusters of sounds, and sequences of speech sounds in words to facilitate meaning, mood, and/or interest (e.g.: in dialogue, monologue, etc.).	1	1	1	1
4. use phonetic and structural generalizations to pronounce unknown words (e.g.: vowel/consonant generalizations, etc.).	1	1	1	1
B. Words (Meaning and Structure)				
1. use conventional word forms (e.g.: verb forms, noun forms, etc.).	2	2	1	1
2. use dialectic word forms when required to facilitate meaning, mood, and/or interest (e.g.: for use in dialogue, monologue, etc.).	1	1	1	1
3. use a variety of types of words (e.g.: vivid words and expressions, precise words, etc.) to meet the requirements of clarity, precision, extensiveness, and stylistic appropriateness.	3	3	3	3
4. use figurative devices to facilitate meaning, mood, and/or interest.	1	1	3	3

---

[1] Conventional is defined here as the pronunciation agreed upon by people who live in the same society. It is well-known and well-accepted pronunciation.

[2] Dialectic is defined here as the variant pronunciation caused by regional and individual speech patterns.

			Speaking Test Objectives (Continued)	Grade Level Priority			
				3	6	9	12
			Skills (The student will be able to . . .)				
4.	B.	5.	develop connotation appropriately in specific contexts to facilitate meaning, mood, and/or interest.	1	2	2	2
		6.	use words and expressions appropriate to the selected language level.[1] (See Organizational Factors, B. - Plan Analysis.)	1	2	3	3
	C.		Sentences (Meaning and Structure)				
		1.	use a variety of sentence (grammatical) patterns to facilitate meaning, mood, and/or interest (e.g.: s-v-o, etc.).	3	3	3	3
		2.	use a variety of sentence levels to facilitate meaning, mood, and/or interest (e.g.: complete/incomplete, simple, compound, complex).	2	2	3	3
		3.	use various sentence types appropriately to facilitate meaning, mood, and/or interest (e.g.: statements, questions, etc.).	3	3	2	1
		4.	rearrange sentences where useful for focus and emphasis.	1	1	2	3
		5.	use a variety of transitional devices.	2	2	3	3
		6.	use a variety of stylistic devices to facilitate meaning, mood, and/or interest (e.g.: repetition of sentence parts, repetition of words, parallelism, etc.).	1	1	3	3
		7.	use sentence patterns, types, and techniques appropriate to the selected language level. See Organizational Factors, B. - Plan Analysis.)	1	2	3	3

---

[1] Language levels may be classified as casual, informal, and formal. Each level has specific features and is used for a specific audience, purpose, situation, and format in an oral communication interaction. It should be noted that these features vary in oral and written communication.

Speaking Test Objectives (Continued)	Grade Level Priority 3	6	9	12
Skills (The student will be able to . . .)				
5. Organizational Factors				
A. Ideational Fluency				
1. produce associations in the form of ideas, opinions, feelings, attitudes, etc., based on a given stimulus.	3	3	3	3
2. repress associations which are irrelevant, confusing, or distracting based on a given stimulus.	1	1	1	1
B. Plan Analysis				
1. select the topic.	2	2	3	3
2. narrow the topic in terms of preparation and presentation time.	1	3	3	3
3. collect and record information and associations about that topic.	1	3	3	3
4. select an appropriate speaking role (e.g.: personal/official; narrator/spectator; etc.).	1	1	1	2
5. select a point of view and keep it consistent.	1	1	1	2
6. select an appropriate language level.	1	2	3	3
C. Organization of Details				
1. state the main idea of a set of details in topic sentence form.	3	3	3	3
2. vary the placement of the topic sentence to facilitate meaning, mood, and/or interest.	1	1	1	1
3. use only details that directly support the main idea.	2	2	3	3
4. use appropriate types of supporting details (e.g.: statistics, expert opinions, illustrations, anecdotes, examples, analogies, etc.) to facilitate meaning, mood, and/or interest.	1	1	2	3

Speaking Test Objectives (Continued)	Grade Level Priority			
	3	6	9	12
Skills (The student will be able to . . .)				
5. C. 5. use appropriate supporting materials (e.g.: print, audio-visual, objects, etc.).	1	1	1	1
6. develop a beginning, middle, and ending (introduction, body and conclusion).	3	3	3	3
7. vary techniques for developing beginnings, middles, and endings.	1	1	1	2
8. use a variety of transitional devices for clarity and emphasis.	1	2	3	3
9. sequence details according to an appropriate organizational pattern (e.g.: time, place, classification, etc.).	3	3	3	3
10. provide details that reveal conflict in attitudes, feelings, opinions, ideas, etc.	1	1	2	3
11. develop setting by providing when and where details.	1	1	1	1
12. develop characterization by providing "inner" and "outer" details.	1	1	1	1
13. develop plot by providing details that indicate conflicts, complications, suspense, and lead to a climax and outcome.	1	1	1	1
14. provide a summary.	1	1	1	1
15. provide a paraphrase.	1	1	1	1
16. develop a written or mental outline.	1	1	1	1

6. Evaluation Factors[1]

A. judge the impact of environmental sound factors. (See #1 - A for objectives and priorities).

---

[1] Evaluation here involves a speaker judging his/her own performance and prioritization is related to that of objectives listed earlier.

Speaking Test Objectives (Continued)	Grade Level Priority			
	3	6	9	12
Skills (The student will be able to . . .)				
6. B. evaluate a communication in terms of voice production factors. (See #2 A-D for objectives and priorities.)				
C. evaluate a communication in terms of nonverbal factors. (See #3 A-D for objectives and priorities.)				
D. evaluate a communication in terms of linguistic factors. (See #4 A-C for objectives and priorities.)				
E. evaluate a communication in terms of organizational factors. (See #5 A-C for objectives and priorities.)				

Appendix A-4

Initial Listening/Speaking Test Objectives: Attitudes

Listening/Speaking Test Objectives	Grade Level Priority[1]			
	3	6	9	12
Attitudes				
1. Motivation (The student will . . .)				
A. experience a level of arousal[2] which will facilitate a communication interaction.	3	3	3	3
B. derive appropriate[3] satisfaction from a communication interaction.	2	2	2	2
C. exhibit a willingness to be tentative (avoid closure).	1	1	1	1
D. exhibit a willingness to focus on a topic and seek clarification about it even if it is not considered particularly important or interesting.	1	1	1	1
2. Image of self and others (The student will . . .)				
A. possess a positive self-image of him/herself as an inter- or intra-communicator.	3	3	3	3
B. exhibit a willingness to become aware of the needs,[4] ideas, feelings, and point of view of others in a communication interaction.	2	2	2	2
C. exhibit a willingness to use a communication interaction to extend his/her awareness of his/her own communication strengths and weaknesses.	1	1	1	1

---

[1] 3 = high priority; 2 = some priority; 1 = minimal priority.

[2] _Arousal_ is defined here as the extent to which the brain is activated as a result of a stimuli. It can be both facilitating and debilitating, as in the case of communication anxiety.

[3] _Appropriate_ is defined here as whatever the communicator requires to take care of his/her needs during a communication situation.

[4] _Needs_ is defined here as a drive to do or achieve something.

Listening/Speaking Test Objectives (Continued)	Grade Level Priority			
	3	6	9	12
Attitudes				
3. Willingness to communicate (The student will exhibit a willingness to . . .)				
A. participate in a communication situation.	1	1	1	1
B. use the social rules appropriate in communication interactions (e.g.: one person talks at a time; speakers take turns; listeners focus attention on the topic and the speaker; frequent and lengthy pauses are avoided; etc.).	1	1	1	1
C. communicate for different reasons and in different ways (e.g.: different purposes, modes, audiences, situations, etc.).	2	2	2	2
D. engage in critical feedback (criticism) both self and other-directed.	1	1	2	2
E. receive critical feedback (criticism) both self and other-directed.	1	1	1	1
F. make changes where appropriate based on critical feedback.	1	1	1	1
G. accept the regional and social dialects of others.	1	1	1	1
H. accept a variety of appropriate roles.	2	2	2	2

Appendix B-1

Initial, Pilot, Field, and Final Listening Test Objectives

Grades 3, 6, 9, 12

Domain	Objective: The student will be able to or know how to . . .	Objective Number				Grade Level-Final Tests			
		Initial	Pilot	Field	Final	3	6	9	12
1 (Pilot) Environmental Sound Factors	match environmental sounds with their venironmental referrents.	1A1	1A1	---	---	-	-	-	-
	distinguish sounds that are appropriate within a context from those which are inappropriate or distracting.	1A2	1A2	---	---	-	-	-	-
	match envoronmental sounds with associated feelings, emotions, and/or attitudes.	1B3	1B3	---	---	-	-	-	-
2 (Pilot) 1 (Field; Final) Voice Production Factors	(Pitch) relate intonation to meaning, mood, and/or interest	2B1	2B1	2B1	1-1	X	X	X	X
	(Quality) relate vocal quality and/or tone of voice to meaning, mood, and/or interest.	2D1	2D1	2D1	1-2	X	X	X	X
	(Rate and Fluency) relate pace to meaning, mood, and/or interest.	2E1	2E1	2E1	1-3	X	X	-	-
3 (Pilot) 2 (Field; Final) Linguistic Factors: Words	understand the meanings of denotative words and phrases in specific contexts.	4B1	4B1	4B1	2-1	X	X	X	X
	understand connotative meanings of words in specific contexts.	4B3	4B3	4B3	2-2	-	X	X	X
	understand a variety of types of words and phrases as they are to facilitate meaning, mood, and/or interest.	4B7	4B7	4B7	2-3	X	X	X	X

Appendix B-1 (Continued)

Domain	Objective: The student will be able to or know how to . . .	Objective Number Initial	Pilot	Field	Final	Grade Level-Final Tests 3	6	9	12
4 (Pilot) 3 (Field; Final) Linguistic Factors: Sentences	understnad messages presented in various sentence types, orders, etc.	4C2	4C2	4C2	3-1	-	-	X	X
	understand transitional devices used to clarify and emphasize meaning and mood.	4C4	4C4	---	---	-	-	-	-
	identify the purpose of stylistic devices.	4C5	4C5	4C5	3-2	-	-	X	X
5 (Pilot) (Not in Field; Final) Organizational Factors: Plan Analysis	indicate the selected topic.	5B1	5B1	---	---	-	-	-	-
	indicate the speaker's purpose in terms of audience and situation.	5B2	5B2	---	---	-	-	-	-
	indicate the selected language level.	5B5	5B5	---	---	-	-	-	-
6 (Pilot) 4 (Field; Final) Organizational Factors: Literal Comprehension	indicate the main idea or main impression of a set of details.	5C1A	5C1A	5C1A	4-1	X	X	X	X
	indicate the details that directly support an idea.	5C1C	5C1C	5C1C	4-2	X	X	X	X
	indicate different types of details that support main idea.	5C1D	5C1D	5C1D	4-3	-	-	X	X
	identify beginning, middle, and ending as well as sequence of details	5C1E	5C1E	5C1E	4-4	X	X	X	X
	identify and understand the use of transitional devices to clarify meaning and mood.	5C1G	5C1G	5C1G	4-5	-	-	X	X
	identify type of organizing pattern and details used to produce pattern.	5C1H	5C1H	5C1H	4-6	X	X	X	X
	indicate the selected language level.	---	---	5C1P	4-7	-	X	X	X

Appendix B-1 (Continued)

Domain	Objective: The student will be able to or know how to . . .	Objective Number Initial	Pilot	Field	Final	Grade Level-Final Tests 3	6	9	12
7 (Pilot) 5 (Field; Final) Organizational Factors: Interpretive Comprehension	infer the thoughts, feelings, emotions of others including the speaker.	5C2C	5C2C	5C2C	5-1	X	X	X	X
	draw conclusions in terms of the intent of the communication.	5C2D	5C2D	---	---	-	-	-	-
	suggest alternative hypotheses and course of action.	5C2F	5C2F	5C2F	5-2	X	X	X	X
	infer the speaker's purpose in terms of the communication.	5C2H	5C2H	5C2H	5-3	X	X	X	X
	infer the speaker's perceived role in terms of the communication.	5C2K	5C2K	---	---	-	-	-	-
	infer overall mood and tone.	5C2O	5C2O	---	---	-	-	-	-
8 (Pilot) 6 (Field; Final) Organizational Factors: Critical Comprehension	judge whether the content of the talk met the speaker's purpose.	5C3A	5C3A	---	---	-	-	-	-
	judge the effect on the communication of the speaker's choice of role, point of view, and/or language level.	5C3B	5C3B	5C3B	6-1	-	-	X	X
	judge if the message is factual or fictitious.	5C3D	5C3D	5C3D	6-2	X	X	-	-
	judge the appropriateness of supporting types of details in terms of message effect.	5C3E	5C3E	5C3E	6-3	-	-	X	X
	judge the relevancy of details.	5C3H	5C3H	---	---	-	-	-	-
	evaluate speaker bias and/or prejudice.	5C3J	5C3J	---	---	-	-	-	-

Appendix B-1 (Continued)

Domain	Objective: The student will be able to or know how to . . .	Objective Number				Grade Level-Final Tests			
		Initial	Pilot	Field	Final	3	6	9	12
8 (Pilot) 6 (Field; Final) Organizational Factors: Critical Comprehension	judge the wisdom and consequences of a charactor's actions in terms of the information presented.	5C3M	5C3M	---	---	-	-	-	-
	draw conclusions in terms of the intent of the communication.	---	---	5C30	6-4	X	X	X	X

Final Test Objectives (N)

Grade	Number
3	14
6	16
9	20
12	20
Total (All Tests)	22

Appendix B-2

Initial, Pilot, Field, and Final Speaking Test Objectives

Grades 3, 6, 9, 12

Domain	Objective: The student will be able to or know how to . . .	Objective Number				Grade Level-Final Tests			
		Initial	Pilot	Field	Final	3	6	9	12
Environmental Sounds	match environmental sounds with their environmental referrents.	K-1-1	---	---	---	-	-	-	-
	match environmental sounds with associated feelings, emotions and/or attitudes.	K-1-2	---	---	---	-	-	-	-
1 Voice Production Factors	project the voice appropriately.	S2A1	1A1	1A1	1-1	X	X	X	X
	use vocal stress to facilitate meaning, mood, and/or interest.	S2A2	1A2	1A2	1-2	X	-	-	-
	indicate point in a communication which should be stressed to facilitate meaning, mood, and/or interest.	K2-3	1A3	1A3	1-3	X	X	X	-
	indicate ways in which voice production factors (volume, pitch, rate and quality) affect the meaning and mood of the message.	K2-1	1A4	1A4	1-4	-	X	-	X
	use intonation to facilitate meaning, mood, and/or interest.	S2B1	1B1	1B1	1-5	-	-	X	-
	use pleasant and natural voice.	S2D1	1C1	1C1	1-6	X	X	X	X
	vary vocal quality to facilitate meaning, mood, and/or interest.	S2D2	1C2	1C2	1-7	-	X	X	X
	indicate where to pause in order to facilitate meaning, mood, and/or interest.	K2-1	1D1	1D1	1-8	X	X	X	X
	pause to facilitate meaning, mood, and/or interest.	S2C1	1D2	1D2	1-9	X	X	X	X

Appendix B-2 (Continued)

Domain	Objective: The student will be able to or know how to . . .	Objective Number				Grade Level-Final Tests			
		Initial	Pilot	Field	Final	3	6	9	12
1 Voice Production Factors	indicate ways in which rate affects meaning, mood, and/or interest.	K2-1	1E1	1E1	1-10	-	X	X	X
	communicate without hesitations, revisions and fillers.	S2E3	1E2	1E2	1-11	X	X	X	X
	use appropriate pace.	S2E1	1E3	1E3	1-12	X	X	X	X
	speak for a prescribed length of time. (1)	---	1E5	1E4	1-13	X	X	X	X
	vary pace to facilitate meaning, mood, and/or tone.	S2E1	1E6	1E5	1-14	-	-	X	X
2 Non-verbal Factors	use posture and gesture appropriately.	S3A1 S3B1	2A1	2A1	2-1	-	-	-	-
	avoid distracting movements, etc.	S3B2	2A2	2A2	2-2	-	-	X	X
	maintain eye contact with audience.	S3C1	2B1	2B1	2-3	X	X	X	-
	use appropriate bodily movements and facial expressions to facilitate meaning, mood, and/or interest.	S3B2 S3D1	2B2	2B2	2-4	X	X	X	X
	determine at which points in a communication a speaker should look at his audience.	K3-3	2B3	2B3	2-5				
3 Linguistic Factors: Sounds	articulate speech sounds clearly and precisely.	S4A1	3A1	3A1	3-1	X	X	X	X
	use conventional pronunciation. (2)	S4A2	3A2	3A2	3-2	X	X	X	X

---

[1] Added to determine length of time student is able to speak.

[2] 3-1 and 3-2 combined at Grade 12 level.

Appendix B-2 (Continued)

Domain	Objective: The student will be able to or know how to . . .	Objective Number Initial	Pilot	Field	Final	Grade Final Level-Tests 3	6	9	12
4 Linguistic Factors-Words	use conventional word forms (e.g.: verb forms, noun and pronoun forms, tense, etc.). (3)	S4B1 K4.2B2	3B1	3B1	4-1	-	-	-	-
	use a variety of types of words (vivid, precise).	S4B3	3B2	3B2	4-2	X	X	X	X
	use stylistic devices.	S4B4	3B3	2. (3C5)	---	-	-	-	-
	use words and expressions appropriate to selected language level.	S4B6 K4.1-6	SB4	sB3	4-3	-	-	X	X
	use oral vocabulary consistent with age norms.	K4.2B1	3B5	3B4	4-4	-	-	-	-
	understand denotative meanings of words and phrases in context.	K4.2B6	3B6	3B5	4-5	X	X	X	-
	understand connotative meanings of words and phrases in context.	S4B5 K4.2B6	3B7	3B6	4-6	X	X	X	X
	use a variety of connectors.(2)	K4.2B5	3B8	(BC4)	---	-	-	-	-
5 Linguistic Factors-Sentences	use a variety of sentence patterns.	S4C1 K4.2C1	3C1	3C1	5-1	-	-	-	-
	use a variety of sentence levels to facilitate meaning, mood, and/or interest.	S4C2 K4.2C1	3C2	3C2	5-2	-	X	X	X
	use a variety of sentence types.	S4C3 K4.2C1	3C3	3C3	5.3	-	-	-	-
	use a variety of transitional devices.	S4C5 K4.2C3	3C4	3C4	5-4	-	X	X	X
	use a variety of stylistic devices.	S4C6 K4.2C4	3C5	3C5	5-5	-	-	X	-

---

[2] 3-1 and 3-2 combined at Grade 12 level.

[3] Combined with sentence domain (5).

Appendix B-2 (Continued)

Domain	Objective: The student will be able to or know how to . . .	Objective Number				Grade Level-Final Tests			
		Initial	Pilot	Field	Final	3	6	9	12
5 Linguistic Factors-Sentences	match language level to audience and occasion.	S4C7	3C6	---	---	-	-	-	-
	use a variety of stylistic devices which facilitate meaning, mood, and/or interest.	S4C4 K4.2C4 K4.2C5	3C7	3C6	5-6	-	-	X	X
	use syntax correctly. (3)		3C8	3C7	5-7	X	X	X	X
	use complete sentences. (3)		3C9	3C8	5-8	X	X	X	X
	present ideas precisely. (3)		3C10	---	---	-	-	-	-
6 Organizational Factors-Ideational Fluency	produce associations in the form of ideas.	S5A1 K5A1	4A1	4A1	6-1	X	X	-	-
	repress associations which are irrelevant, confusing or distracting.	S5A2	4A2	4A2	6-2	-	-	-	-
7 Organizational Factors-Plan Analysis	select and/or narrow a topic.	K5B1 K5B2 S5B1 S5B2	4B1	4B1	7-1	-	-	-	-
	select and organize information.	K5B3 S5B3	4B2	4B2	7-2	-	-	-	-
	state the purpose of a talk.	S5C1	4B3	4B3	7-3	-	-	X	X
	develop an outline for a talk.	K5C19 S5C16	4B4	4B4	7-4	-	-	-	-
	infer meaning, mood or tone. (4) (6)		4B5	4B5	7-5	-	X	X	X
	communicate effectively (overall). (7)		4B6	---	---	-	-	-	-

---

[3] Combined with sentence domain (5). [4] Combined with sentence domain (5^1).

[5] Specific extension of previous objectives.

[6] Left out unintentionally on original objectives.

[7] Attempt to evaluate total communication rather than underlying skills.

Appendix B-2 (Continued)

Domain	Objective: The student will be able to or know how to . . .	Objective Number Initial	Pilot	Field	Final	Grade Level-Final Tests 3	6	9	12
8 Organizational Factors-Details	state main idea of a set of details.	K5C1 S5C1	4C1	4C1	8-1	X	-	X	X
	exclude details that do not support main idea.	S5C3	4C2	4C2	8-2	X	X	X	X
	select appropriate types of supporting details.	K5C6 K5C7 S5C4	4C3	4C3	8-3	-	-	-	X
	develop a beginning, middle, and ending or develop an introduction, body and conclusion.	K5C8 S5C7	4C4	4C4	8-4	X	X	X	X
	sequence details according to an appropriate organizational pattern (time, place, classification, etc.).	K5C9 K5C10 S5C9	4C5	4C5	8-5	X	X	X	X
	draw inferences from a set of details.	K5C17	4C6	4C6	8-6	X	X	X	-
	summarize details.	K5C18 S5C14	4C7	4C7	8-7	-	-	-	-
	emphasize important points.	K5C2 K5C3	4C8	4C8	8-8	-	-	-	-
	select and develop a topic sentence.	K5C5	4C9	4C9	8-9	-	-	X	X

Final Test Objectives (N)

Grade	Number
3	23
6	28
9	35
12	32
Total (All Tests)	118

Appendix C-1

Final Listening Test Objectives, Grade 3

Domain and Number	Objective: Student knows how to or is able to . . .	Obj. #	# of Items
1 Voice Production Factors	relate intonation to meaning, mood, and/or interest.	1-1	2
	relate vocal quality and/or tone of voice to meaning, mood, and/or interest.	1-2	5
	relate pace to meaning, mood and/or interest.	1-3	2
2 Linguistic Factors - Words	understand the meanings of denotative words and phrases in specific contexts.	2-1	3
	understand a variety of types of words and phrases as they are used to facilitate meaning, mood, and/or interest.	2-3	2
3 Linguistic Factors - Sentences		0	0
4 Organizational Factors - Literal Comprehension	indicate the main idea or main impression of a set of details.	4-1	2
	indicate the details that directly support an idea.	4-2	2
	identify beginning, middle, and ending as well as sequence of details.	4-4	3
	identify type of organizing pattern and details used to produce pattern.	4-6	2
5 Organizational Factors - Interpretive Comprehension	infer the thoughts, feelings, emotions of others including the speaker.	5-1	2
	suggest alternative hypotheses and courses of action.	5-2	3
	infer the speaker's purpose in terms of the communication.	5-3	2
6 Organizational Factors - Critical Comprehension	judge if the message is factual or fictitious.	6-2	2
	draw conclusions in terms of the intent of the communication.	6-4	3

Objective Total - 14 Item Total - 35

Appendix C-2

Final Listening Test Objectives, Grade 6

Domain and Number	Objective: Student knows how to or is able to . . .	Obj. #	# of Items
1 Voice Production Factors	relate intonation to meaning, mood, and/or interest.	1-1	4
	relate vocal quality and/or tone of voice to meaning, mood, and/or interest.	1-2	2
	relate pace to meaning, mood, and/or interest.	1-3	2
2 Linguistic Factors - Words	understand the meanings of denotative words and phrases in specific contexts.	2-1	3
	understand connotative meanings of words in specific contexts.	2-2	2
	understand a variety of types of words and phrases as they are used to facilitate meaning, mood, and/or interest.	2-3	2
3 Linguistic Factors - Sentences		0	0
4 Organizational Factors - Literal Comprehension	indicate the main idea or main impression of a set of details.	4-1	2
	indicate the details that directly support an idea.	4-2	2
	identify beginning, middle, and ending as well as sequence of details.	4-4	3
	infer the speaker's purpose in terms of the communication.	4-6	2
	indicate the selected language level.	4-7	3
5 Organizational Factors - Interpretive Comprehension	infer thoughts, feelings, emotions of others including the speaker.	5-1	2
	suggest alternative hypotheses and courses of action.	5-2	3
	infer the speaker's purpose in terms of the communication.	5-3	2

Appendix C-2 (Continued)

Domain and Number	Objective: Student knows how to or is able to . . .	Obj. #	# of Items
6 Organizational Factors - Critical Comprehension	judge if the message is factual or fictitious.	6-2	2
	draw conclusions in terms of the intent of the communication.	6-4	4

Objective Total - 16 Item Total - 40

Appendix C-3

Final Listening Test Objectives, Grade 9

Domain and Number	Objective: Student knows how to or is able to . . .	Obj. #	# of Items
1 Voice Production Factors	relate intonation to meaning, mood, and/or interest.	1-1	2
	relate vocal quality and/or tone of meaning, mood, and/or interest.	1-3	3
2 Linguistic Factors-Words	understand the meanings of denotative words and phrases in specific contexts.	2-1	3
	understand connotative meanings of words in specific contexts.	2-2	2
	understand a variety of types of words and phrases as they are used to facilitate meaning, mood, and/or interest.	2-3	3
3 Linguistic Factors-Sentences	understand messages presented in various sentences types, orders, etc.	3-1	2
	identify the purpose of stylistic devices to facilitate meaning, mood, and/or interest.	3-2	2
4 Organizational Factors-Literal Comprehension	indicate the main idea or main impression of a set of details.	4-1	3
	indicate the details that directly support an idea.	4-2	3
	indicate different types of details that support main idea.	4-3	2
	identify beginning, middle, and ending as well as sequence of details.	4-4	2
	identify and understand the use of transitional devices to clarify meaning and mood.	4-5	4
	identify type of organizing pattern and details used to produce pattern.	4-6	2
	indicate the selected language level.	4-7	2

Appendix C-3 (Continued)

Domain and Number	Objective: Student knows how to or is able to . . .	Obj. #	# of Items
5 Organizational Factors-Interpretive Comprehension	infers the thoughts, feelings, emotions of others including the speaker.	5-1	3
	suggest alternative hypotheses and courses of action.	5-2	2
	infer the speaker's purpose in terms of the communication.	5-3	2
6 Organizational Factors-Critical	judge the effect of the speaker's choice of role, point of view, and/or language level on the communication.	6-1	2
	judge the relevance and appropriateness of types of supporting details in terms of message effect.	6-3	2
	draw conclusions in terms of the intent of the communication.	6-4	4

Objective Total - 20 Item Total - 50

Appendix C-4

Final Listening Test Objectives, Grade 12

Domain and Number	Objective: Student knows how to or is able to . . .	Obj. #	# of Items
1 Voice Production Factors	relate intonation to meaning, mood, and/or interest.	1-1	2
	relate vocal quality and/or tone of voice to meaning, mood, and/or interest.	1-3	2
2 Linguistic Factors-Words	understand the meanings of denotative words and phrases in specific contexts.	2-1	2
	understand connotative meanings of words in specific contexts.	2-2	2
	understand a variety of types of words and phrases as they are used to facilitate meaning, mood, and/or interest.	2-3	3
3 Linguistic Factors-Sentences	understand messages presented in various sentence types, orders, etc.	3-1	2
	identify the purpose of stylistic devices to facilitate meaning, mood, and/or interest.	3-2	2
4 Organizational Factors-Literal Comprehension	indicate the main idea or main impression of a set of details.	4-1	2
	indicate the details that directly support an idea.	4-2	2
	indicate different types of details that support main idea.	4-3	3
	identify beginning, middle, and ending as well as sequence of details.	4-4	2
	identify and understand the use of transitional devices to clarify meaning and mood.	4-5	3
	identify type of organizing pattern and details used to produce pattern.	4-6	1
	indicate the selected language level.	4-7	4

Appendix C-4 (Continued)

Domain and Number	Objective: Student knows how to or is able to . . .	Obj. #	# of Items
5 Organizational Factors-Interpretive Comprehension	infer the thoughts, feelings, emotions of others including the speaker.	5-1	3
	suggest alternative hypotheses and courses of action.	5-2	3
	infer the speaker's purpose in terms of the communication.	5-3	3
6 Organizational Factors-Critical Comprehension	judge the effect of the speaker's choice of role, point of view, and/or language level on the communication.	6-1	2
	judge the relevance and appropriateness of types or supporting of details in terms of message effect.	6-3	3
	draw conclusion in terms of the intent of the communication.	6-4	4

Objective Total - 20 Item Total - 50

Appendix C-5

Final Speaking Test Objectives, Grade 3

Domain and Number	Objective: Student knows how to or is able to . . .	Obj. #	# of Items
1 Voice Production Factors	project the voice appropriately.	1-1	1
	use vocal stress to facilitate meaning, mood, and/or interest.	1-2	2
	use a pleasant and natural voice.	1-6	1
	indicate where to pause to facilitate meaning, mood, and/or interest.	1-8	1
	pause to facilitate meaning, mood, and/or interest.	1-9	1
	communicate without hesitations, or fillers.	1-11	2
	use appropriate pace.	1-12	2
	speak for a prescribed length of time.	1-13	1
2 Nonverbal Factors	maintain eye contact with audience.	2-3	1
	use appropriate bodily movements and facial expressions to facilitate meaning, mood, and/or interest.	2-4	2
3 Linguistic Factors-Sounds	articulate speech sounds clearly and precisely.	3-1	3
	use conventional pronunciation.	3-2	1
4 Linguistic Factors-Words	use a variety of types of words (vivid, precise).	4-2	2
	understand denotative meanings of words in context.	4-5	2
	understand connotative meanings of words and phrases in context.	4-6	1
5 Linguistic Factors-Sentences	use syntax correctly.	5-7	3
	use complete sentences.	5-8	1
6 Organizational Factors-Ideational Fluency	produce associations in the form of ideas.	6-1	1
7 Organizational Factors-Plan Analysis		---	-

Appendix C-5 (Continued)

Domain and Number	Objective: Student knows how to or is able to . . .	Obj. #	# of Items
8 Organizational Factors-Details	state the main idea of a set of details.	8-1	2
	exclude details that do not support main idea.	8-2	1
	develop a beginning, middle, and ending.	8-4	1
	sequence details according to an appropriate organizational pattern, e.g.: time, place, classification, etc.	8-5	4
	draw inferences from a set of details.	8-6	2

Objective Total - 23 Item Total - 38

Appendix C-6

Final Speaking Test Objectives, Grade 6

Domain and Number	Objective: Student knows how to or is able to . . .	Obj. #	# of Items
1 Voice Production Factors	project the voice appropriately.	1-1	1
	indicate point in a communication which should be stressed to facilitate meaning, mood, and/or interest.	1-3	2
	indicate ways in which voice production factors (volume, pitch, rate and quality) affect the meaning and mood of the message.	1-4	1
	use pleasant and natural voice.	1-6	1
	vary vocal quality to facilitate meaning, mood, and/or interest.	1-7	2
	indicate where to pause in order to facilitate meaning, mood, and/or interest.	1-8	2
	pause to facilitate meaning, mood, and/or interest.	1-9	1
	indicate ways in which rate affects meaning, mood, and/or interest.	1-10	2
	communicate without hesitations, revisions and fillers.	1-11	2
	use appropriate pace.	1-12	2
	speak for a prescribed time period.	1-13	1
2 Nonverbal Factors	maintain eye contact with audience.	2-3	1
	use appropriate body movements and facial expression to facilitate meaning, mood, and/or interest.	2-4	2
3 Linguistic Factors-Sounds	articulate speech sounds clearly and precisely.	3-1	2
	use conventional pronunciations.	3-2	1
4 Linguistic Factors-Words	use a variety of types of words (vivid, precise).	4-2	2
	understand denotative meanings of words and phrases in context.	4-5	2
	understand connotative meanings of words and phrases in context.	4-6	2

Appendix C-6 (Continued)

Domain and Number	Objective: Student knows how to or is able to . . .	Obj. #	# of Items
5 Linguistic Factors-Sentences	use a variety of sentence levels to facilitate meaning, mood, and/or interest.	5-2	1
	use a variety of transitional devices.	5-4	1
	use syntax correctly.	5-7	3
	use complete sentences.	5-8	1
6 Organizational Factors-Ideational Fluency	produce associations in the form of ideas.	6-1	1
7 Organizational Factors-Plan Analysis	infer meaning, mood, or tone.	7-5	1
8 Organizational Factors-Details	exclude details that don't support main idea.	8-2	1
	develop a beginning, middle, and ending.	8-4	1
	sequence details according to an appropriate organizational pattern (time, place, classification, etc.).	8-5	2
	draw inferences from a set of details	8-6	2

Objective Total - 28 Item Total - 43

Appendix C-7

Final Speaking Test Objectives, Grade 9

Domain and Number	Objective: Student knows how to or is able to . . .	Obj. #	# of Items
1 Voice Production Factors	project voice appropriately.	1-1	1
	indicate point in communication which should be stressed to facilitate meaning, mood, and/or interest.	1-3	2
	use intonation to facilitate meaning, mood, and/or interest.	1-5	1
	use a pleasant, natural voice.	1-6	1
	vary vocal quality to facilitate meaning, mood, and/or interest.	1-7	2
	indicate where to pause in order to facilitate meaning, mood, and/or interest.	1-8	2
	pause to facilitate meaning, mood, and/or interest.	1-9	1
	indicate ways in which rate affects meaning, mood, and/or interest.	1-10	2
	communicate without hesitations, revisions or fillers.	1-11	2
	use appropriate pace.	1-12	2
	speak for prescribed length of time.	1-13	1
	vary pace to facilitate meaning, mood, and/or tone.	1-14	1
2 Nonverbal Factors	avoid distracting movements, etc.	2-2	1
	maintain eye contact with audience.	2-3	1
	use appropriate body movements and facial expressions to facilitate meaning, mood, and/or interest.	2-4	2
3 Linguistic Factors - Sounds	articulate speech sounds clearly and precisely.	3-1	1
	use conventional pronunciation.	3-2	1
4 Linguistic Factors - Words	use a variety of types of words (vivid, precise).	4-2	1
	use words and expressions appropriate to selected language level.	4-3	1

3[illegible]9

Appendix C-7 (Continued)

Domain and Number	Objective: Student knows how to or is able to . . .	Obj. #	# of Items
4 Linguistic Factors - Words	understand the denotative meanings of words and phrases in context.	4-5	1
	understand connotative meanings of words and phrases in context.	4-6	2
5 Linguistic Factors - Sentences	use a variety of sentence levels to facilitate meaning, mood, and/or interest.	5-2	1
	use a variety of transitional devices.	5-4	1
	use a variety of stylistic devices.	5-5	3
	use a variety of stylistic devices which facilitate meaning, mood, and/or interest.	5-6	1
	use syntax correctly.	5-7	2
	use complete sentences.	5-8	1
6 Organizational Factors - Ideational Fluency		---	-
7 Organizational Factors - Plan Analysis	state purpose of talk.	7-3	1
	infer meaning, mood, and/or tone.		3
8 Organizational Factors - Details	state main idea of a set of details.	8-1	1
	exclude details that do not support main idea.	8-2	1
	develop an introduction, body, and conclusion.	8-4	4
	sequence details according to an appropriate organizational pattern (time, place, classification, etc.).	8-5	3
	draw inferences from a set of details.	8-6	1
	select and develop a topic sentence.	8-9	2

Objective Total - 35 Item Total - 54

Appendix C-3

Final Speaking Test Objectives, Grade 12

Domain and Number	Objective: Student knows how to or is able to . . .	Obj. #	# of Items
1 Voice Production Factors	project voice appropriately.	1-1	1
	indicate ways in which voice production factors (volume, pitch, rate and quality) affect the meaning and mood of the message.	1-4	2
	use a pleasant, natural voice.	1-6	2
	vary vocal quality to facilitate meaning, mood, and/or interest.	1-7	1
	indicate where to pause in order to facilitate meaning, mood, and/or interest.	1-8	2
	pause to facilitate meaning, mood, and/or interest.	1-9	1
	indicate ways in which rate affects meaning, mood, and/or interest.	1-10	1
	communicate without hesitations, revisions or fillers.	1-11	2
	use appropriate pace.	1-12	2
	speak for prescribed length of time.	1-13	1
	vary pace to facilitate meaning, mood, and/or tone.	1-14	1
2 Nonverbal Factors	avoid distracting movements, etc.	2-2	1
	use appropriate body movements and facial expressions to facilitate meaning, mood, and/or interest.	2-4	2
	determine at which points in a communication a speaker should look at his audience.	2-5	1
3 Linguistic Factors - Sounds	articulate speech sounds clearly and precisely.	3-1 + 3-2 Combined	2
	use conventional pronunciations.		
4 Linguistic Factors - Words	use a variety of types of words (vivid, precise).	4-2	1

Appendix C-8 (Continued)

Domain and Number	Objective: Student knows how to or is able to . . .	Obj. #	# of Items
4 Linguistic Factors - Words	use words and expressions appropriate to selected language level.	4-3	1
	understand connotative meanings of words and phrases in context.	4-6	3
5 Linguistic Factors - Sentences	use a variety of sentence levels to facilitate meaning, mood, and/or interest.	5-2	1
	use a variety of transitional devices.	5-4	1
	use a variety of stylistic devices which facilitate meaning, mood, and/or interest.	5-6	3
	use syntax correctly.	5-7	2
	use complete sentences.	5-8	1
6 Organizational Factors - Ideational Fluency		6-1	-
		6-2	-
7 Organizational Factors - Plan Analysis	state purpose of talk.	7-3	1
	infer meaning, mood, and/or tone.	7-5	2
8 Organizational Factors - Details	state main idea of a set of details.	8-1	1
	exclude details that do not support main idea.	8-2	1
	select appropriate types of supporting materials.	8-3	1
	develop an introduction, body, and conclusion.	8-4	4
	sequence details according to an appropriate organizational pattern (time, place, classification, etc.)	8-5	3
	select and develop a topic sentence.	8-9	3

Objective Total - 32 Item Total - 51

Appendix D-1

Final Listening Test Objectives/Items Matrix, Grade 3

Domain	Objective	Obj. #	Subtest 1	2	3	4	5
1	Voice Production Factors	1-1					30, 31
		1-2				23	32, 33 34, 35
		1-3			11	21	
2	Linguistic Factors - Words	2-1			14	16, 19	
		2-2					
		2-3				18, 22	
3	Linguistic Factors - Sentences	3-1					
		3-2					
4	Organizational Factors - Literal Comprehension	4-1		5	12		
		4-2	4		15		
		4-3					
		4-4		6, 7		17	
		4-5					
		4-6		9		28	
						28	
5	Organizational Factors - Interpretive Comprehension	5-1	1, 2				
		5-2	3	8		25	
		5-3			13	26	
6	Organizational Factors - Critical Comprehension	6-1					
		6-2		10		27	
		6-3					
		6-4				20, 24 29	

Total Number of Objectives = 14 Total Number of Items = 35

Appendix D-2

Final Listening Test Objectives/Items Matrix, Grade 6

Domain	Objective	Obj. #	Subtest 1	2	3	4	5	6	7
1	Voice Production Factors	1-1				19		31, 36	37
		1-2			13			29	
		1-3		10				30	
2	Linguistic Factors - Words	2-1		6				27	39
		2-2	1		15				
		2-3		11					38
3	Linguistic Factors - Sentences	3-1							
		3-2							
4	Organizational Factors-Literal Comprehension	4-1			12		21		
		4-2					22	28	
		4-3							
		4-4			16		23, 24		
		4-5							
		4-6					25		40
		4-7				17, 18			
5	Organizational Factors-Interpretive Comprehension	5-1	3	8					
		5-2	2	7				33	
		5-3			14			35	
6	Organizational Factors-Critical Comprehension	6-1							
		6-2	4	9					
		6-3							
		6-4	5				26	32, 34	

Total Number of Objectives = 16 Total Number of Items = 40

Appendix D-3

Final Listening Test Objectives/Items Matrix, Grade 9

Domain	Objective	Obj. #	Subtest 1	2	3	4	5	6	7	8
1	Voice Production Factors	1-1				25			41	
		1-2						35	43,44	
		1-3								
2	Linguistic Factors - Words	2-1	1							47,48
		2-2		5					38	
		2-3	2	6	18					
3	Linguistic Factors - Sentences	3-1		7						50
		3-2	3,4							
4	Organizational Factors-Literal Comprehension	4-1			14		27			45
		4-2		10			28		42	
		4-3		9,13						
		4-4					32,33			
		4-5			16,17		30,31			
		4-6					29			49
		4-7				23,24				
5	Organizational Factors-Interpretive Comprehension	5-1			22			34		46
		5-2		11	21					
		5-3		8	15					
6	Organizational Factors-Critical Comprehension	6-1		12					39	
		6-2								
		6-3			20		26			
		6-4			19			36,37	40	

Total Number of Objectives = 20 Total Number of Items = 50

Appendix D-4

Final Listening Test Objectives/Items Matrix, Grade 12

Domain	Objective	Obj. #	Subtest 1	2	3	4	5	6	7	8	9
1	Voice Production Factors	1-1							40	42	
		1-2						35		43	
		1-3									
2	Linguistic Factors - Words	2-1									48,49
		2-2		5							46
		2-3		6	18						47
3	Linguistic Factors - Sentences	3-1		7						45	
		3-2	3								50
4	Organizational Factors-Literal Comprehension	4-1	1				27				
		4-2		10			29				
		4-3		9,13			30				
		4-4					28			44	
		4-5			16,17		33				
		4-6				23,24	32				
		4-7			15	25					
5	Organizational Factors-Interpretive Comprehension	5-1	4		22			34			
		5-2		11	21				38		
		5-3	2	8	14						
6	Organizational Factors-Critical Comprehension	6-1		12					41		
		6-2									
		6-3			20		31,26				
		6-4			19			36,37	39		

Total Number of Objectives = 20 Total Number of Items = 50

Appendix D-5

Final Speaking Test Objective/Items Matrix, Grade 3

Domain	Objective	Obj. #	Subtest			
			1	2	3	4
1	Voice Production Factors	1-1				29
		1-2			11	31
		1-6				30
		1-8		10		
		1-9				32
		1-11			27	37
		1-12			12	33
		1-13			28	
2	Nonverbal Factors	2-3			13	
		2-4			14,24	
3	Linguistic Factors-Sounds	3-1			19	35
		3-1				36
4	Linguistic Factors-Words	4-2			15,20	
		4-5	3,4			
		4-6	2			
5	Linguistic Factors-Sentences	5-7			22,23	
		5-8			21,24	
6	Organizational Factors-Ideational Fluency	6-1			25	
7	Organizational Factors-Plan Analysis	---				
8	Organizational Factors-Details	8-1		6,9		
		8-2			17	
		8-4			18	
		8-5		7,8		
		8-6	1,5			

Appendix D-6

Final Speaking Test Objective/Items Matrix, Grade 6

Domain	Objective	Obj. #	Subtest 1	2	3	4
1	Voice Production Factors	1-1				36
		1-3	5,9			
		1-4		10		
		1-6				38,37
		1-7			16	39
		1-8	7	14		
		1-9				40
		1-10	8	11		
		1-11				34
		1-12			17	
		1-13				35
2	Nonverbal Factors	2-3			18	
		2-4			19	42
3	Linguistic Factors-Sounds	3-1			25	43
		3-2				44
4	Linguistic Factors-Words	4-2			20,26	
		4-5	2	12		
		4-6	3,4			
5	Linguistic Factors-Sentences	5-2			21	
		5-4			27	
		5-7			28,29,30	
		5-8			31	
6	Organizational Factors-Ideational Fluency	6-1			32	
7	Organizational Factors-Plan Analysis	7-5	1	15		
8	Organizational Factors-Details	8-2			23	
		8-4			24	
		8-5			22,33	
		8-6	6	13		

Appendix D-7

Final Speaking Test Objectives/Items Matrix, Grade 9

Domain	Objective	Obj. #	Subtest 1	2	3	4	5
1	Voice Production Factors	1-1					46
		1-3	2,7				
		1-5	6				
		1-6				24	47
		1-7					48
		1-8	4,5				
		1-9					49
		1-10	3,9				
		1-11				41	54
		1-12				25	51
		1-13				42	
		1-14					50
2	Nonverbal Factors	2-2				27	
		2-5	8				
		2-4				28	52
3	Linguistic Factors-Sounds	3-1				26	53
4	Linguistic Factors-Words	4-2				36	
		4-3				34	
		4-5		12			
		4-6	6	13,14			
5	Linguistic Factors-Sentences	5-2				29	
		5-4				39	
		5-5				43,44 45	
		5-6		11			
		5-7				37,38	
		5-8				30	
6	Organizational Factors-Ideational Fluency	---					

Appendix D-7 (Continued)

Domain	Objective	Obj. #	Subtest 1	2	3	4	5
7	Organizational Factors-Plan Analysis	7-3				35	
		7-6	1	10,15			
8	Organizational Factors-Details	8-1			17		
		8-2				31	
		8-4		18	19	32,33	
		8-5			22,23	40	
		8-6		16			
		8-9			20,21		

Appendix D-8

Final Speaking Test Objectives/Items Matrix, Grade 12

Domain	Objective	Obj. #	Subtest			
			1	2	3	4
1	Voice Production Factors	1-1				43
		1-4	3,11			
		1-6			20	44
		1-7				45
		1-8	1,4			
		1-9				46
		1-10	2			
		1-11			38	51
		1-12			21	48
		1-13			39	
		1-14				47
2	Nonverbal Factors	2-2			23	
		2-4			24	49
		2-5	10			
3	Linguistic Factors-Sounds	3-1			22	50
		3-2				
4	Linguistic Factors-Words	4-2			32	
		4-3			30	
		4-6	7,8,9			
5	Linguistic Factors-Sentences	5-2			25	
		5-4			35	
		5-5			40,41 42	
		5-7			33,34	
		5-8			26	
6	Organizational Factors-Ideational Fluency	---				
7	Organizational Factors-Plan Analysis	7-3			31	
		7-5	5,6			

Appendix D-8 (Continued)

Domain	Objective	Obj. #	Subtest 1	 2	 3	 4
8	Organizational Factors-Details	8-1		12		
		8-2			27	
		8-3			36	
		8-4		13,14	28,29	
		8-5		18,19	37	
		8-9		15,16 17		